Ethics and the emotions : an account of the role of sentiment in moral thought

Grant, Stephen Michael

Ethics and the Emotions: An Account of the Role of Sentiment in Moral Thought

Stephen Grant
King's College London, 2005

PhD

Acknowledgements

During my time at King's College London I had the good fortune of being jointly supervised by Raimond Gaita and Peter Goldie, both of whom have had an enormous influence on the development of my ideas, and to whom I am deeply grateful. James Hill also provided very helpful comments on Chapter Six. An earlier version of Chapter Four was delivered to the Manchester Metropolitan University Social Sciences Research Seminar, and parts of Chapters Two and Three were delivered as a talk to the King's College London Senior Research Seminar. I am grateful for the useful comments I received in each case. I have also benefited greatly from numerous discussions of much of the material contained here with my colleagues at Richmond upon Thames College, Paul Sheehy, Paul Sperring and Adrian Samuel, and I am particularly grateful to Paul Sheehy for several helpful suggestions concerning the chapter on practical necessity.

Table of Contents

In short, moral codes too are only a *sign language of emotions.*

Nietzsche

Introduction

The relationship between morality and the emotions is a theme which has provoked debate within ethics for perhaps as long as it has existed as an identifiable area of study. Yet in spite of its recurring presence across different periods of philosophy, and despite the great disparities in the style and content of thinkers who have written on the subject, there is a surprising simplicity to the way in which philosophers have tended to view the emotions in this context. They are generally seen either as a powerful and dangerous distraction to the operation of reason - the true basis for morality, or as being at the very heart of morality, making it possible for us to be the kind of moral beings we are, and indispensable to the leading of a moral life.

The two philosophers who are perhaps most emblematic in this sense are Kant and Hume. For Kant, 'inclination, be it good-natured or otherwise, is blind and slavish'.[1] The emotions were incapable of generating maxims in accordance with the moral law, and could at best provide additional (but nevertheless unwelcome) support for the effectiveness of the moral law in motivating us. To be free entailed being free from the influence of inclination, and being governed by pure practical reason. Hume, by contrast, famously argued that morality was 'more properly felt than judg'd of'.[2] The sentiment of benevolence, as well as our wider sensual capacities, were absolutely central to morality, and reason could do no more than guide the passions.

The divisions are not always this stark, but the same sort of concerns arise constantly with regard to how we should understand the role of the emotions. Plato identified 'spiritedness' as one of the three elements in his tripartite division of the soul, and argued that it is the 'natural auxiliary of the rational, so long as it is not corrupted by evil upbringing.'[3] Aristotle agreed that having the right education of the emotions was central to becoming a virtuous agent, and saw the cultivation of the right sort of feelings as a crucial step along the path to becoming a virtuous agent.[4] Once again, the image is of emotion having a potentially overwhelming force which needs to be guided if we are to be capable of right action. This sort of debate also captures much of the content of contemporary moral philosophy when it deals with the issue of the emotions, with

---

[1] Immanuel Kant, *Critique of Practical Reason*, 124.
[2] David Hume, *Treatise of Human Nature*, 470.
[3] Plato, *Republic*, 441a.
[4] See in particular Aristotle, *Nicomachean Ethics*, Books I and II.

distinctly Kantian thinkers standing opposed to those whose allegiance is closer to Hume or Aristotle.[5]

The aim of this thesis will be to defend a form of sentimentalism. More specifically, I shall argue that compassion and reason are jointly necessary and sufficient for a proper morality. A study of compassion will be the focus of chapter six, but it is worth taking some time now to indicate what I intend to show. By compassion, I shall mean an emotion which involves the perception that someone is suffering or deprived, coupled with an associated feeling, and a desire to alleviate the suffering or deprivation. In claiming that reason must be involved, I mean that it is only when reason comes to be both influenced by compassion, as well as guiding our compassionate responses, that we can acquire a full moral outlook. An infant or some animals may have the capacity to act in what we might describe as a compassionate manner, but they could not understand it as compassion, or be motivated to act because they see a particular action as the compassionate thing to do. It is these features or the virtuous agent which are made possible by the onset of reason.

Talk of a 'proper' morality reflects the view that there could be acts which fall under a description employing moral concepts, but which might not be possible to interpret in ways which are consistent with compassion. To take an obvious example, the heroic, noble type who has entered what Nietzsche described as the 'extra moral' phase,[6] will act in ways which reflect a set of moral standards, but not ones which involve the Christian compassion of sickly, modern European man. If there can be morality without compassion, then compassion cannot account for all morality. In the face of this, I shall claim that compassion does account for any proper morality, and this claim will be cashed out in the context of a naturalistic psychology. By this I mean that given the kind of beings we are, ones capable of suffering and responding to the suffering of others, capable of happiness and promoting the happiness of others, then a morality based on compassion and guided by reason is proper in virtue of its reflecting the kind of creatures we are.

In order to defend this line of argument, I shall structure the thesis into two distinct sections. The first three chapters will be an enquiry into the nature of emotion, and the following five chapters will be a consideration of the relationship of the emotions to our

---

[5] Amongst the legion of works which deal with this issue, prominent, recent contributions of a distinctly Humean or Aristotelian line which defend a role for the emotions in ethics include Simon Blackburn, *Ruling Passions*, Alan Gibbard, *Wise Choices, Apt Feelings*, Philippa Foot, *Natural Goodness*, Rosalind Hursthouse, *On Virtue Ethics*. For positions which are more distinctly Kantian, see Thomas Nagel, *The Possibility of Altruism*, and Christine Korsgaard, *The Sources of Normativity* and *Creating the Kingdom of Ends*.

[6] Friedrich Nietzsche, *Beyond Good and Evil*, § 32.

ethical lives, with a particular emphasis on compassion. There are two reasons for structuring the work in this way. The first is that if there is going to be an attack or a defence of the role of emotion within ethics, then it is best to be clear about the nature of emotion. Such a policy makes less likely the danger of advancing certain premises which rely on an inaccurate characterisation of one of the key areas of debate. In order to highlight this, let me offer an example of where such a case has occurred in the past literature.

A J Ayer's treatment of morality was to claim it was composed of a series of 'pseudo-concepts'. He claimed that when advancing supposedly ethical judgements, 'the function of the relevant ethical word is purely "emotive". It is used to express feelings about certain objects, but not to make any assertion about them.'[7] Given that such language is calculated purely to excite emotions in others, then there could be no real agreement or disagreement in ethics, because the moral part of the statements didn't constitute a statement about the world - it was simply an expression of one's own feelings about the world. It is an unexamined premise of Ayer's argument that the expression of feeling cannot be the subject of agreement or disagreement, and this in turn suggests a theory of the emotions as simple, unanalysable phenomena, akin to sensations, which include no judgements about the world. If this account of emotion and feeling is wrong, then the wider moral theory may be dependent on a false premise, and I suggest that this is exactly what occurred with emotivism. The theory of emotions which I shall be defending is to the effect that they are complex phenomena which involve beliefs about the world, and which can therefore be the subject of disagreement. If for example, someone is afraid of the neighbour's dog on the grounds that she believes all dogs bite humans at every opportunity, the fear (I shall argue) involves the belief that she is in danger from all dogs, and there is plenty of room for discussion and disagreement over whether or not her fear is appropriate in this context. By failing to consider adequately the nature of emotion, Ayer assumes that if moral language is founded in our moral lives then it cannot be a source of agreement or disagreement, and this now looks like a questionable premise.

The second principal reason for addressing in some detail the nature of emotion is that I believe it will provide important insights into the nature of morality, and once again, let me offer an example of how this might work. One of the claims I shall make is that although emotions are naturally occurring in humans, they are 'developmentally open'. That is to say that we are born with a capacity for fear, for example, but precisely what we actually come to fear will be partially dependent on environmental circumstances. This means that those

born into different communities may well come to share quite general fears, such as fear of physical harm, but there may also be dramatic discontinuities between what some of us come to fear and some don't. To take an extreme example, contrast the attitude of the London stockbroker and the African farmer with regard to falling share prices. The first may be so terrified by the prospect of this that she becomes suicidal, whereas the second may look on with complete indifference. Yet each may react in a similarly fearful way towards a car approaching them at high speed. As such, it looks as if the content of our emotional lives is less consistent across communities than physical traits such as visual perception, but more consistent than language. This conclusion has considerable implications if we are to defend a moral theory which has a key role for the emotions. It indicates that if this sort of pattern is typical, and our moral lives are in some way dependent upon our emotional development, we may expect a divergence between the content of differing sets of moral judgements. Furthermore, if the content of our emotional and moral lives is partly constituted by our locatedness within a particular community, there may be points of divergence between different communities which are irresolvable.

A further feature of our emotions which contributes to the claim that they are complex phenomena is the role of desires. It is widely argued in the contemporary literature that emotions can be understood as beliefs or judgements. I shall be supporting the view that all emotions involve beliefs or perceptions, which is how they acquire the property of intentionality - the state of being directed towards the world. But in arguing that emotions necessarily involve feelings I will be confronting some of these contemporary theories. A further move away from such 'cognitivist' views will come in the form of the relationship which I shall try to show exists between our emotions and desires. In some cases, desires are necessarily involved in emotions, most obviously in instances of envy and jealousy. If I envy my neighbour his new car, a desire for such a car or that particular car, or of her status in light of the possessing of the car is a necessary constituent of the emotion. Similarly, if I am jealous of my colleague's relationship with our boss, then I desire such a relationship, or else I desire that she should not have it. In such cases, the desire is bound up with the beliefs and feelings, all of which are jointly involved in the emotion.

One way in which the role of desire adds a further twist to the complexity of our emotions is that one can argue that they are necessary constituents of emotions such as those mentioned above, but only contingently part of certain other emotions. I might feel a burning pride at a victory in a tennis match and, whilst in the throes of the emotion, invite

---

[7] A J Ayer, *Language, Truth and Logic*, 111.

my opponent to play again. One way of construing this is that the desire to issue the invitation was part of the general emotional of experience of pride at the winning of the match. That said, it is equally plausible that I might have been satisfied at the victory, and felt no desire to play again, or any other desire to act which could be explained in light of the emotion.

These issues point in turn to the rather tricky issue of motivation. It is, I think, uncontroversial to claim that emotions have motivational potential, but if we are presented with theories of the emotions which range from cognitivist accounts on the one hand which attempt to reduce emotions to beliefs or judgements, through to complex 'feeling-laden' accounts on the other which claim that emotions contain beliefs or perceptions, feelings, and sometimes desires, then we need to be clear about what provides emotions with the motivational force emotions they have. The account offered here will be broadly consistent with a belief/desire account of motivation, claiming that emotions have motivational force either because they involve desires, or because they cause desires to emerge. In practice this might mean that our envy directly involves a desire for an object, and that desire explains our later action of purchasing the object. Or it might be the case that I am flushed with pride, and later recollection of the pleasing nature of the pride provokes a desire to repeat the experience which was the source of the pride in the first place. As such, the appearance of the desire is explained by a thought along the lines of 'I want to experience that emotion again'.

I have said that the thesis is divided into two broad sections, but I shall now offer a more detailed account of the line of argument I shall be taking. In the initial chapters on the nature of emotion, there are two key themes which will emerge. The first is an opposition to reductivism, and in particular to any attempt to reduce emotions either to the status of simple sensations, or to an attitude towards a proposition. Both sorts of reduction, I shall claim, result in an inadequate account of what emotions are, and as such would have a distorting effect on any later moral theory. In the case of the first sort of reduction, that of portraying emotions as purely internal physiological phenomena, the danger will be that the later moral theory will suffer from the sort of problem encountered in emotivism. If, on the other hand, we try the second form of reduction, that of analysing emotions into a series of beliefs, it may lead perversely to a sentimentalist theory of ethics which is also cognitivist. If ethics is founded ultimately on, for example, compassion,[8] and compassion can be analysed into a series of beliefs, then we could accept the first premise, carry out the reduction, and establish whether or not the beliefs which go to make up compassionate

responses in a particular instance are true or false representations of the world. As such, the validity of moral judgements need not differ greatly from other putative empirical claims about the world.

At a purely intuitive level, the idea of a cognitivist, sentimentalist theory of ethics just seems profoundly odd, and I shall attempt to rule out such a possibility in arguing that emotions are complex phenomena which include minimally either beliefs or perceptions on the one hand, and feelings on the other, which is to say both cognitive and non-cognitive elements.[9] As such, any reductive programme is doomed to failure if it attempts to reduce emotion to one or other of these elements. Moreover, as I shall be arguing that these are necessary features of emotion, the dismissal of reduction will be an *a priori* claim.

Part of the complexity of any emotion derives from the necessary presence of feelings, which I shall characterise as perceptions of the state of one's body, and this will lead directly to the second key claim I shall be making in the first section. I shall show that there is a 'background of directed feeling' which orients us towards the world, and which is shaped by our social environment. This means that through innumerable transactions with others around us over many years, we come to associate certain feelings with the world around us. These feelings condition our perception of many features of the world, and guide us in our dealings with other persons. To offer an example, when a child comes across an unfamiliar object, and the caregiver looks on with horror as the child approaches it, this response is one to which a child often responds, and is capable of deterring her from going near it. This sort of simple example is symptomatic of the way in which we come to 'feel towards' the world in one way rather than another. It is when this sort of scenario occurs on many occasions with regard to a wide variety of persons, objects and activities, that we come to 'see' the world in some ways rather than others, in ways which are conditioned by the way in which we feel about it.

The opening chapters will develop this theory at some length, but represent primarily the groundwork for the ethical theory which is to follow. The bridge between the theory of the emotions in the initial three chapters and the moral theory of the final five chapters will be a discussion of what it is to have a moral perspective, which I shall characterise as a part of our emotional development. The focus of the discussion will be a concept prominent

---

[8] This is Schopenhauer's claim in *On the Basis of Morality*.

[9] I do not claim that emotions are characterised only be beliefs and feelings, in that there are other features such as their evolutionary development or their neural correlates which one might argue to be essential. For this reason, I say only that emotions necessarily *include* beliefs and feelings, as it is these features which are most relevant to the later moral theory.

of Bernard Williams' discussion in *Shame and Necessity* - that of the internalised other. I shall characterise this is a metaphor for the way in which our emotional responses come to be shaped by those around us, resulting in the needs and claims of others coming to be a part of our own identity. This means that when we confront the world, we do so from a perspective which already includes the presence of others' interest within it, thereby creating the possibility of altruism.[10]

The next stage of the thesis will be to justify the claim that there is a sort of content that the internalised other ought to have in light of the kind of beings we are, but in making this claim I shall consider in some detail what I take to be a major hurdle in reaching such a conclusion. Let us say we make the claim, as Schopenhauer did, that morality is based in compassion, and that the meaning of moral terms can be understood in terms of compassion. If this is true, then it should be true *a priori* that compassionate responses are good responses, but this just doesn't seem to be the case. We might even argue, with Nietzsche, that this smacks of Christian pity, and that we should rid ourselves of such sickly modern European thoughts. As such, we arrive at how the emotion binds us to the world, but not yet how it relates to morality. The way to overcome such a conceptual hurdle is by means of a naturalistic psychology which allows us to conclude that the content of the internalised other is one which should be reflective of our nature as beings who both need those around us, and have the capacity to respond to such needs in others. Compassion is a response which characterises both our needs and such possible responses, and this will be the basis for the claim that it grounds a proper morality.

The final chapters will confront two particularly challenging problems for any sentimentalist theory of ethics, those of practical necessity and moral reasoning. Given that emotions are generally taken to be a dangerous distraction to sound reasoning, the sentimentalist is under pressure to provide an account of whether or not such reasoning could take place, and if so, how. As with, the discussion of emotion and morality in the preceding chapters, I shall argue that a sound understanding of the nature of emotion, one which I hope to have provided in the opening chapters, will provide satisfactory answers on these issues. Once we have a view of emotions as partly constituted by beliefs or perceptions, and that these are coloured by feelings, then insofar as our beliefs and perceptions are similarly 'feeling-laden',[11] we have the potential for agreement and disagreement. For example, if we each hold the feeling-laden belief that human life has

---

[10] This is, needless to say, a radically different explanation to that provided by Nagel in his book of the same name.

[11] I take this phrase from Michael Stocker. See in particular his *Valuing Emotions*.

great value, but disagree over issues related to euthanasia, we have the potential for a resolution of our disagreement if one of us can bring forward certain evidence which will convince the other that the value of human life is better respected by one side rather than the other in this debate. The shared feelings which are combined in consciousness with a similar range of beliefs held by each agent provide a platform upon which discussion can be conducted in much the same way as ordinary debate - drawing inferences, weighing the evidence, reaching conclusions. In contrast, should we not share the feeling-laden belief that animals' lives have value, then such debate would be impossible, but there is, so I shall argue, sufficient communality of interest amongst humans such that broadly similar emotional and feeling-laden responses to the world are more rather than less likely to result in the possibility of moral reasoning and discussion.

Prior to this, I shall consider how it is that sentimentalism can account for our sense of moral obligation. The difficulty here is addressing the Kantian claim that the categorical nature of moral obligation requires that we ground it in the moral law. If we are to claim that obligation is based in contingent features of the human condition such as compassion, then all imperatives become hypothetical ones along the lines that 'if I wish to be compassionate, then I need to help those in need'. The danger is that this may not be adequate to account of the force of our sense of obligation, and sentimentalism therefore encounters the danger of being unable to account for a key area of moral thought.

The line I shall take in defence of how sentimentalism can indeed explain our sense of obligation stems from arguments of one finds in Hume and in Bernard Williams.[12] Hume provides us with an account of how a sense of obligation can emerge from a combination of self-interest and our sentiment of benevolence, and I shall argue in favour of his account of how we come to accept a 'general rule' on the grounds that we recognise that the rules of justice benefit us personally. This is supplemented by a sense of fellow-feeling with others, such that we come to dislike those who flout justice even when we personally do not stand to lose. This, I shall claim, is consistent with the development of a moral sense which can be captured in the concept of the internalised other. Our emotions are shaped such that we come to feel a particular way about the world, and this includes acquiring a respect for rules of justice. It is Williams' discussion of what he calls 'dispositions of character' which provide the grounding of our sense of obligation. That is to say, the compassionate agent is one who perceives the world in light of her

[12] The key sections in Hume come in Book III of *A Treatise of Human Nature*, and Williams' key arguments are to be found in 'Practical Necessity' and 'Moral Incapacity', reprinted in his *Moral Luck* and *Making Sense of Humanity* respectively.

compassionate disposition, and her perceptions of the world condition her understanding of the suffering of others, and motivate her to act on perceiving that suffering. There are some occasions when our reasons for action overwhelm any other reasons we may have to act differently, and this is when we feel we *must* act in one way rather than another.

There are two further comments it is important to make at this stage. The first is that the overwhelming emphasis will be on advancing a positive theory of how the emotions help ground morality. In particular, I aim to provide a plausible account of how the emotions contribute to our moral psychology. This means that I shall not be concerned to attack opponents of sentimentalism, although the final two chapters are intended to provide answers in the areas of moral philosophy where sentimentalism is traditionally considered to be most vulnerable. More specifically, I shall try to confront what I take to be the most pressing concerns that a Kantian is likely to have, those in the areas of the nature of moral obligation and the process of moral reasoning. I shall not on the other hand be setting out the positive case that the Kantian would make with the aim of attacking it. The second point is that I shall not be addressing the question of how sentimentalism relates to the content of our concept of justice. Although I would like to have considered this issue, considerations of time and space made it impractical adequately to address this in addition to writing on both the nature of emotion, and how emotion relates to morality.

The overall thrust of this thesis will be that it is through our emotions that the world takes on significance for us, and that we not only encounter a world which is enlivened by the emotional responses it elicits in us, but find ourselves bound to others by those same emotional commitments. The wide range of emotions we experience in a plethora of different contexts impels us into an engagement with the world, and part of that engagement comes in the form of responding to the hold that others have over us through our emotional ties to them. It is here where we find the moral domain.

1. Cognitivism in the Emotions

Introduction

I said in the main introduction that part of what I wish to argue for is the view that our emotional life has a central role to play in our engagement with the world. Our emotions help us to find colour and wonder in the world around us, and bind us to certain persons and projects. One aspect of this for which I shall be arguing is that there is a distinct ontology of the emotions, and this claim would be vulnerable if it could be shown that emotions, far from being distinct, can be reduced to a combination of beliefs, or beliefs and desires. This has been a view widely held within analytical philosophy from the second half of the twentieth century onwards, and in order to fortify the position I shall be trying to establish, it is necessary to consider the essentials of such 'cognitivist' views. I shall look more closely at the role of desire in the following two chapters, but my aim here will be to consider positions which defend an account of emotions as collections of beliefs.

There are four characteristics, most or all of which one typically finds in the positions I shall be examining.

- emotions can be understood purely in terms of the beliefs they entail
- feelings are not part of an emotion
- emotions have propositional content
- emotions have intentionality

I shall argue that the first two claims are wrong, but that the second two are correct, and provide us with a genuine advance in our understanding of emotion. My aim is to show that we need to avoid any reductive account which singles out either beliefs or feelings as being the defining feature of emotions, and accept an account of emotions as complex phenomena comprising both beliefs and feelings. As such, the position I shall refer to as cognitivism would be wrong in the narrowness of its focus, but right in placing an emphasis on the role of belief.

It is perhaps wise in the first instance to underline how radical the shift towards cognitivism was, in that emotions had traditionally been considered as simple, unanalysable states, largely physiological in character.[1] To see them discussed as beliefs flies in the face not simply of many traditional historical approaches, but also of intuitive thinking about emotions as being feelings which defy full analysis or understanding. We just are afraid or angry, and there isn't anything more which can be said about the content of such states. Although I shall try to show that congnitivist accounts of the emotions cannot succeed in the final analysis, I shall also be emphasising the great leap forward which such accounts represent in comparison to the historical ones they have countered. I shall argue that attempting to understand emotions as beliefs and desires is fatally flawed, but that even those philosophers who have roundly rejected such theories have largely accepted key aspects of the cognitivist position[2], and the fact that contemporary accounts often stress the complexity of the issue indicates the influence of many of the accounts I shall be discussing.

The James-Lange Theory

I have said that the aim of this chapter is principally to discuss cognitivist accounts of the emotions, but I shall be starting with one of the classic non-cognitivist accounts, as it has provided so much ammunition for the cognitivists, and represents something of an historical starting point for contemporary philosophy of the emotions. The James-Lange theory essentially sees emotions as perceptions we have of states within our own bodies.[3] James consciously inverts the conventional framework for understanding our emotions.

> Our natural way of thinking about these standard emotions is that the mental perception of some fact excites the mental affection called the emotion, and that this latter state of mind gives rise to the bodily expression. My thesis on the contrary is that *the bodily changes follow directly the* PERCEPTION *of the exciting fact, and that our feeling of the same changes as they occur* IS *the emotion.*[4]

---

[1] For an historical account of theories of the emotions in ancient, medieval and early modern philosophy, see Susan James, *Passion and Action.*
[2] See for example Michael Stocker, *Valuing Emotions*, 38-51, Peter Goldie, *The Emotions*, 11-28.
[3] See William James, 'What is an Emotion?', *Mind 9*, 188-205. Carl Lange was a Danish psychologist credited by James with having developed the same theory independently of him at around the same time.
[4] William James, 'What is an Emotion?', *Mind 9*, 190-191.

In other words, rather than a perception giving rise to bodily changes which we take to be part of fear, anger, disgust, and so on, the body responds directly to an external stimulus, and our perception of our racing heart or trembling knees is the emotion. This is the legacy of our evolutionary design, in that we have developed in such a way that the brain is adapted to respond in a particular way when presented with certain circumstances. A number of arguments are invoked to support the theory, of which some are of particular importance when we come to consider the cognitivist views which I shall be looking at next. James argues that one piece of evidence supporting his thesis is that if we try to abstract away all the feelings and the sensations which are characteristic of strong emotions such as great fear or anger, then all we are left with is a rather flat perceptual state. Or to use James' more colourful terms, 'What kind of an emotion of fear would be left, if the feelings neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is quite impossible to think.'[5]

There are certain aspects of James' theory which are of central importance for what follows. The first is that James lays enormous emphasis on the role played by feelings in his thesis. Throughout the article we are treated to numerous passages similar to the one quoted above in which the presence of feeling is seen as the defining characteristic of those emotions with 'distinct bodily expression' in which he is interested. It is this aspect of his work which was to be rejected by the cognitivists, largely perhaps, because it was felt to be closely associated with what is universally regarded by the overwhelming weakness of the theory. The problem is that if an emotion is the perception of a bodily state, then how can we distinguish between emotions which have identical bodily states? If my heart beats quickly, this can be out of fear, anger, excitement, elation or a range of other emotions. We could try turning to combinations of behavioural activity and physiological activity to narrow down the range of options, so that we might look at the combination of running, screaming, accelerated heart rate, sweating and certain other features to define fear. But this would only pick out an individual emotion in a small number of cases, and only a minority even of those emotions with 'bodily expression' in which James is interested.

---

[5] Ibid. 193-194

One obvious direction in which to turn is towards the object of our emotions. If the emotion is one of fear, and the object of the emotion is an elephant charging towards us, then our beliefs about the approaching elephant provide a means of distinguishing this particular emotion from others. Unfortunately for James, this option was denied him, because the object of the emotion was a bodily state rather than an external object. As such, we can't distinguish very easily which emotion is instantiated on a particular occasion by my heart thumping or my muscles tightening. Although there is 'an exciting fact', this isn't the object of the emotion on a Jamesian account. One way of understanding the problem with James' account, is that he has restricted the complexity of his conceptual account. By describing emotions as perceptions of bodily states, he provides them with intentionality, but the intentionality does not provide the right sort of relationship to the external world. We want an account which explains our fear with direct reference to the charging elephant, and this relationship is not explicit enough in the account we have before us.

The problem of providing too narrow an ontology of the emotions is precisely the same one encountered by the cognitivists, who overcome James' problem over how emotions bring us into contact with the external world, but leave out James' most powerful observations as to how they affect us internally.

The Intentionality of the Emotions

Perhaps the most prominent cognitivist account of emotions has been offered by Anthony Kenny, whose principal focus was precisely on the intentionality of emotions.[6] To talk of intentionality in this context means that our emotions must always be directed towards someone or something - they must have an object. For Kenny, not only must any emotion have an object, but 'emotions are specified by their objects'.[7] This is established in light of two different features of objects in language.[8] In the first place, there is the logical role of object-expressions related to a particular class of *verbs* in which objects describe species of the genus described by the verb. For example, stealing a car is one specific species of stealing, and smoking a cigarette is a species of smoking. In the second place, objects can also pick out the *formal objects*. The formal object of ϕing is the

---

[6] Anthony Kenny, *Action, Emotion and Will*, especially 187-194
[7] *ibid*, 73

object under the description which must apply if it is possible to ϕ at all. Obvious examples which Kenny cites are that only what is edible can be eaten, and only what is inflammable can be burned. But there are also more informative cases, such as only what is dirty can be cleaned, and only what is wet can be dried.

Given the claim that all emotions must have objects, then the formal requirements on what could possibly constitute the objects of love, fear, envy and so on, will inevitably provide a much broader understanding of our emotions in two senses. Firstly, the intentionality on its own sheds light on the central importance of seeing that emotions cannot be understood as simply sensations or private phenomena. They necessarily force us into a relationship with the world around us. This is an obvious break with the Jamesian account set out earlier, and conforms to the powerful intuition that there is a strong relationship between our emotions and the features of the external world which appear to elicit them. Secondly, the formal restrictions, Kenny argues, will restrict the range possibilities as to when it is possible for us to experience particular emotions. As he puts it, 'One cannot be afraid of just anything, nor happy about anything whatsoever'.[9] Peter Goldie highlights two further necessary properties for an emotion's object. It must be sufficiently fine-grained to explain why the person feels the way he does. 'Oedipus might be delighted that he married Jocasta, but not delighted that he married his mother.' Furthermore, the object need not exist, in that it is perfectly coherent to be afraid of ghosts or the abominable snowman.[10]

The first of Kenny's points does seem to provide a genuine insight into our emotions. We must fear something or love someone, and it is this feature of emotions which allows us to break out of the conceptual restrictions of the body which Kenny detects in earlier philosophers such as Descartes and Hume. This aspect of our emotions is also one which has been largely accepted by those who have followed Kenny, regardless of whether or not they are sympathetic to his cognitivism.[11]

---

[8] *ibid*, 188-189
[9] *ibid*, 192
[10] Peter Goldie, *The Emotions*, 18
[11] Among the many philosophers who accept the intentionality of the emotions, see Robert C Solomon, *The Passions*, or Gabrielle Taylor, *Pride, Shame and Guilt*, for alternative cognitivist accounts, or else see Justin Oakley, *Morality and the Emotions*, Michael Stocker, *Valuing Emotions*, Peter Goldie, *The Emotions*, for theories of the emotions which retain the intentionality but reject the distinctly cognitivist approach.

The debate surrounding the formal restrictions on what we can fear or envy is more difficult, and once again points us in the direction of a complex ontology. It seems right to say that that I cannot be proud of something I find unequivocally bad, or that I must find some value in something I envy, but this tells us very little, and even such relatively modest claims can prove difficult. Kenny gives us a useful account of how emotions can be inappropriate, explaining that if I envy my neighbour's fruit trees the envy would become inappropriate on discovering that the trees were actually growing in my own garden[12]. Clearly, certain emotions just ought to cease in light of learning certain facts, but we also need to be clear about just what our story is going to be when it comes to listing the relevant facts. Take for example Kenny's claim that one can't be afraid of just anything. Ideally, we might want to say that either we are afraid only of what we take to be dangerous, or else the fear is inappropriate. But what if we have a case of someone who is scared of spiders, and is able to offer an account along the lines that the thought of such a repellent creature scuttling across her skin is so appalling that she just doesn't want to have a spider anywhere near her? Such an explanation clearly fails to meet the putative formal requirement of fearing what is dangerous, but we can surely present a strong case to the effect that it can be appropriate to fear what we find disgusting, because we find the experience of disgust so awful. This would mean that the object of our fear might be disgust rather than danger.

Kenny is surely right to claim that we can't be afraid of just anything, but the problem is that we can be appropriately afraid of an awfully lot things which aren't dangerous. One obvious route is to try to fall back on certain natural properties of objects as the stimulus for what we fear, and there is some mileage to be gained from this. As William James pointed out, 'In advance of all experience of elephants no child can but be frightened if he suddenly find one trumpeting and charging upon him.'[13] But such brutally clear circumstances account for only a fraction of the cases in which we are afraid. The fear of death or serious bodily injury is much less common for most of us than more complex instances of fear such as fear of losing a job or fear of embarrassment in a philosophy seminar.

---

[12] The issue of when an emotion is appropriate is one I shall consider at various points, but it is worth giving an example now in order to highlight the intuition behind such an idea. Imagine you are in a restaurant and a diner at the next table flies into rage because the young waitress who laid the table had put the knife and fork on the wrong side of one of the places at the table. Part of our view would be that rage was an inappropriate emotional response in the face of such a trivial mistake, and when I talk of inappropriate emotions, I shall be referring to this sort of case.

There is also a danger that if we try to develop a theory along the lines that fear is an emotional response stimulated by natural properties in the external world, then it can lead us in the wrong direction when we come to interpret the object of fear. If I say that I am frightened of what the Australian fast-bowling attack may do to the English batsmen, then the physical dangers present in an Australian hurling a ball at an Englishman might conceivably be the object of my fear, in which case the physical properties of a cricket ball and the potential injury to a human body will provide us with the right sort of explanation. But a more likely interpretation, one which may be lost if we focus on natural properties, is that the object of my fear is not the prospect of physical injury, but the more subtle dangers of swing and seam bowling, the fear of which must be explained within the cultural context. These more subtle and complex examples are absolutely crucial to the debate on the emotions, because they are so central to what makes human life distinct. Although we can make a case of the young boy in James' scenario experiencing fear of the elephant which can be compared to the elk's fear of a lion, the plausibility of such comparisons fade in proportion to the subtlety and complexity of the emotion we are discussing. More specifically, when distinctly human capacities such as language and self-awareness start to figure in the development and experience of emotions such as remorse or jealousy, then it will be ever more difficult to pick out natural properties in the world in order to explain their occurrence. The size and power of an elephant could provide much of the explanation of why we might fear one charging towards us, but in cases such as fear of the possibility of future failure, then there just may not be any salient features of the physical world which could explain this emotion.

A Syllogistic Approach

One major problem which we are left with then, is how emotions come to be directed towards some objects rather than others. A different tack can be found in Donald Davidson's account of emotion, where the focus moves from intentionality to beliefs.[14] Davidson argues that we can extract from Hume's discussion of the passions in Book II of his *Treatise* a more persuasive theory of the emotions. Following Hume, Davidson

---

[13] William James, 'What is an Emotion?', *Mind 9*, 191

[14] Donald Davidson, 'Hume's Cognitive Theory of Pride', reprinted in *Essays on Actions and Events*, 277-290.

focuses on pride, but then proceeds to discard the aspects of Hume which he finds unfortunate. Pride does not entail any particular feelings or sensations, much less are there ones which constitutes pride. These are simply phenomena which may accompany pride on certain occasions, and we need to set aside such 'psychological atomism' in order to unearth the real insights Hume offers us.

Davidson is particularly interested in what he believes Hume reveals about 'propositional pride', which is to say pride evident in sentences such as, 'She was proud that she had been elected president'. It is argued that a certain reading and reorganisation of the double relation of ideas and impressions will provide us with a much better theory. Davidson summarises Hume as follows:

> The cause of pride is a conjunction of the idea of a house, say, and a quality (beauty). The quality causes the separate and pleasant passion, which under the right conditions causes (by association) the similar pleasant passion of pride. The passion of pride itself always causes the idea of self to appear, and this idea must be related (causally, by association) to the idea of the object (the house) on which the quality is placed.[15]

In short, the first impression is the pleasure caused by the beauty of the house. The first idea is the ownership of the house. The second idea is brought to the fore by the first - it is not simply that the house is owned, but it is owned by *me*, thereby bringing out the idea of the self. These features combine to produce pride, and the pleasure of this emotion is the second impression. We need to hold on to the causal role played by certain beliefs, and the relation of the object of pride to self, but to introduce an attitude of *approval* in order to give us the appropriate structure.

According to Davidson, this can best be captured in the form a practical syllogism such as the following;

> All who have the quality of owning a house are praiseworthy in that respect

---

[15] *Ibid.* 282. The original discussion in Hume can be found in pages 285-290 of *A Treatise of Human Nature*. As Davidson makes it clear that he is interested only in what he believes Hume should have

I have the property of owning a house

Therefore

I am praiseworthy in that respect.

The beliefs about my ownership of the house are coupled with the pro attitude towards people with this property combine to cause pride. Davidson claims that this formulation helps to explain pride in two ways, in that it gives both a causal explanation and the person's reasons for being proud.

At this stage, it is worth considering one of the principal criticisms which has been put against this sort of theory, which focuses on the exclusion of the concept of feeling from the accounts offered by both Kenny and Davidson. The problem with the sort of cognitivist account which we have looked at is that it appears that even if all the relevant beliefs are in place with regard to the appropriate object, there is still no guarantee that there is any emotion. Consider, for example, a situation in which two people are walking down the street, and a deadly snake slithers out of the undergrowth and turns towards them. It is possible for both onlookers to form the beliefs that the snake is deadly, and that they are in danger if they do not immediately stop and allow it to continue on its way without approaching it. We can even include relevant desires such as that of avoiding a fatal bite. Nevertheless, it is still possible that the first might be paralysed with fear while the second waits calmly for the snake to pass so they can continue their conversation. The crucial point seems to be that without certain *feelings*, we cannot be sure that the content of the beliefs, coupled with certain desires, will give us what we require for the fear to be present. If it is possible for two agents to hold all of the relevant beliefs and desires, but for one to be experiencing fear and not the other, then it looks as if beliefs and desires cannot be the whole story when it comes to our emotions.

One useful attempt to defend a cognitivist line against this sort of criticism comes from those who agree with the failings in something like the Davidsonian account, but try to retain the cognitivist theory in a different form. The move they make, correctly in my

---

said, rather than in providing a fully accurate interpretation, I have not questioned the summary he offers.

view, is to emphasise the complexity of our emotions by requiring us to understand them within the context of a wider life. This sort of elaboration is offered by both Gabrielle Taylor and Robert C Solomon, although with slight variations.[16] Taylor begins her account with a criticism levelled specifically at Davidson, before developing a richer theory of her own. In the first instance, she points out that the structure of pride (or any other propositional emotion) needn't take the form set out by Davidson. Someone might, for example, decide to purchase a house with purely prudential considerations in mind, then find herself unexpectedly satisfied with the state of being a homeowner, and finally come to the view that anyone in this state ought to feel worthy of admiration for getting there. As such, the major premise is derived from the minor one, rather than the other way round.

The more far-reaching criticism is one which concerns how we should come to understand why someone is experiencing a particular emotion. Rather than explaining this in terms of similar beliefs held by others, there are times when we need to look much more closely at the life of the particular individual concerned, and to identify the place of the emotion within a much wider framework of beliefs and experiences relating solely to the person concerned. In order to bring out this point, consider the following passage from Guiseppe Tomasi Lampedusa's *The Leopard*.[17] The scene takes place in the decaying Palace of the Salinas, a once great family of Sicilian aristocrats. Princess Concetta, now in her seventies, is receiving Tassoni, the Cardinal of Palermo. He has inadvertently let slip that Tancredi, the man with whom she had been in love many years before, had also loved her, but had been put off by what he believed to be her coldness towards him. This shatters her belief that her father had kept them apart in pursuit of his political aims. She has never married or taken a lover since.

> If Tassoni had told the truth, then the long hours spent in savouring her hatred before her father's picture, her hiding of every photograph of Tancredi so as not to be forced to hate him too, had been stupidity – worse, cruel injustice; and she suffered now at the memory of Tancredi's warm and imploring tone as he had begged his uncle to allow him into that convent; they had been words of love

---

[16] Gabrielle Taylor, *Pride, Shame, and Guilt*, 1-16, Robert C Solomon, *The Passions*, and 'Emotions and Choice', in Amelie O. Rorty, ed. *Explaining Emotions* 251-282.

[17] Taylor uses and excerpt from Joyce's *The Dead* in order to make a similar point. I have substituted a different passage which better suits my purposes.

> towards her, words not understood, routed by her pride which at her harshness had drawn back with their tails between their legs like whipped puppies. From the timeless depth of her being, a black pain came welling to spatter her all over that revelation of the truth. (186)

Later I shall suggest that such passages pose considerable difficulties if one tries to reduce them to a series of beliefs, but many of the points which Taylor raises in setting out such a theory nevertheless extend our understanding of what has gone wrong. She suggests that if we are to take a Davidsonian approach to understanding any of the numerous emotions involved in such an episode, then we will lose sight of the interweaving role of the beliefs and experiences of Concetta at the personal level. To choose but one of the emotions involved here, it would be misleading to try to frame her sadness in the form of a syllogism such as:

All people who have missed out on the opportunity of marrying someone they love are sad insofar as they have missed this opportunity.

Concetta has missed this opportunity.

Therefore

Concetta is sad.

The problem comes in trying to phrase the major premise in such a way that it is not open to obvious counter examples. In this instance, it is entirely possible that Concetta's feelings might have receded over the years to the point where she felt little more than a touch of amusement at this news. Alternatively, her beliefs about the immense happiness of Tancredi's marriage to someone else, and the satisfaction at the solitary life she had led might have left her relieved at the way things had worked out. The problem appears to lie in the fact that our emotions often fail to conform to the law-like structure of the syllogism - the wide variety of possible emotional responses to particular situations makes abstraction hazardous because it calls into question whether or not there really could be general laws of logic to describe our emotional responses.

Such arguments indicate that appreciating why an individual experiences a particular emotion at a particular moment is less successful if we try to abstract from a range examples to achieve a general set of laws, and better understood if we focus on emotions as integrated into a wider life. Concetta's sadness can only be understood when considered alongside her love, her hatred, her sense of injustice, her pride, and countless other aspects of her life which have preceded this moment. There are certainly some cases where we can usefully talk in broader terms about the kind of emotions people are likely to experience in particular circumstances, as I conceded with reference to William James earlier,[18] but such cases will become ever rarer once we leave behind the clear-cut beliefs we are likely to develop in the face of charging elephants, and arrive at the more deeply personalised emotions which spring from our more individual circumstances. At this stage, only knowledge of the agent's more particular history, set against the backdrop of who they are, will give us the explanatory material we need to understand why they are having a particular emotion. Taylor herself concludes that, 'the appeal is no longer to the wholly rational being; it is to the admittedly far less neat and precise notion of what it would be human and natural for a person to feel, given that person's other beliefs and attitudes'.[19]

This emphasis on the complex, interlocking role of emotions within a life is echoed in Robert Solomon's early theory. He argued for emotions not as beliefs, but as judgements. Like Taylor, he was at pains to avoid what he saw as the irrelevant use of feeling in the understanding of the emotions. 'It is at the heart of my argument that "feelings" and physiology and, with qualifications, dispositions to behave, do *not* play an essential role in the constitution of emotions.'[20] Like Taylor, Solomon stressed the complexity of the phenomenon, arguing that, 'an emotion is never a single judgement, but a system of judgements, and although one might well make one or several judgements of the system without having the emotion, my claim is that one cannot make *all* of them and not have the emotion.'[21] All the essential judgements necessarily involve oneself in a scenario which one has created. 'Anger, for example, is to be analysed in

---

[18] See page 16 above.
[19] Gabrielle Taylor, *Pride, Shame, and Guilt*, 14
[20] Robert C Solomon, 'Emotions and Choice', Appendix, 274, in Amelie O Rorty ed. *Explaining Emotions*. Although Solomon uses the term 'judgement', there appears to be no relevant distinction between his use of this term and the use of the term 'belief' which we find in Kenny and Davidson. I have therefore addressed his theory on the assumption that the two terms are interchangeable in this context.
[21] *Ibid*, 275

terms of a quasi-courtroom scenario, in which one takes the role of judge, jury, prosecuting attorney and, on occasion, executioner.....The object of anger is the accused, the crime is an offence, and the overall scenario is one of judgemental self-righteousness.'[22]

According to Solomon, the fact that this doesn't appear to reflect the ordinary experience of anger could be explained by the fact the judgements we make which are emotions are not of the category which are the result of deliberation. They are judgements which are 'undeliberated, unarticulated, and unreflective'.[23] The practicalities of life mean that we simply don't have time to carefully think through all the options we will often have before us, and a mechanism which leads us to flee unthinkingly before possible danger is going to be useful in saving our necks more often than not.

Solomon supplements the description of emotions as judgements by giving an account of the functional role emotions play in helping us through ordinary life. 'Emotions are rational responses to unusual situations.....An emotion is a necessary hasty judgement in response to a difficult situation.'[24] Such a view carries obvious force when we come to emotions such as fear or anger, where they fit into an obvious account which one might wish to provide in the context of an evolutionary theoretical account of the development of the emotions.[25]

Once again, there seem to be invaluable observations offered by Taylor and Solomon, but ones which cannot conceal deeper problems with their accounts. Solomon in particular seems to struggle in trying to square his cognitivist aims with the peculiar phenomenology of the emotions, in that he must try to show how the emotions are rational despite the intuition that they are partly characterised by non-rational elements such as physiological changes and unreflective responses. This is particularly evident in the use of the courtroom as a metaphor for anger. It is surely an exaggeration to think that we set up such a scenario in our minds, even subconsciously or metaphorically. Indeed, the formality and organised structure of a courtroom just looks inappropriate even as a metaphor for the potentially uncontrolled and destructive impulses which anger

---

[22] *Ibid*, 275
[23] Robert C Solomon, *The Passions*, 131
[24] Robert C Solomon, 'Emotions and Choice', 264-265, in Amelie O Rorty ed. *Explaining Emotions*
[25] For a thorough treatment of this area, see Paul Griffiths, *What Emotions Really Are*.

often involves. In addition, to think of emotions as ultimately rational responses to the world around us seems to leave out the rather unhelpful, irrational behaviour often associated with such responses. If I argue with my neighbour over a trivial issue, and some of his comments irritate me such that I find it difficult ever to look at him, and spend years resenting him and throwing my grass cuttings over the fence into his back garden, the ongoing nature of my anger does nothing to help me deal with an unfamiliar situation, still less does it help me get through my life.

These points raise problems for Solomon's theory, but the central argument which both he and Taylor put is that emotions have to be understood as being embedded in a life, and it is only in the context of this life that we can understand why someone has a particular emotion. Such an approach is supplemented by certain arguments which Taylor emphasises in terms of what she sees as the rational intelligibility of certain emotions. For example, if I say that I am afraid because I believe the world will end tomorrow, the content of the belief and the fact that it is *my* belief jointly give a reason for my fearing tomorrow's sunrise. This is similar to the point raised by Kenny with regard to my envying my neighbour's fruit trees, and then discovering they are my own. We can make an emotion more or less intelligible, and decide it is more or less appropriate, by considering the beliefs which contribute to it. This is precisely because beliefs have a clear causal potential in the generation of emotions. If I come to believe that you are about to push me over, this can cause fear, and if I do become afraid we will need to explain why I am experiencing this emotion by referring to that belief.

The problem with all of these arguments is that it is far from clear that any of them make even the slightest dent in the view that we cannot have emotions without feelings. If we return for a moment to the scenario of the two people who happen upon a deadly snake, then Solomon would tell us that they simply cannot both have exactly the same set of relevant judgements. If that were the case then they would both experience fear. I assume that Taylor might suggest that a contextualised understanding of how their beliefs about snakes fit into the wider pattern of their beliefs about the world would allow us to see that the superficial similarity disappears once we realise that these beliefs play a different role from one person to the next. For the person who experiences no fear to believe that he is in no danger is a belief honed by experience to the point where it plays a different role in the way in which it leads one to see the world. It is at this

point where it is time to tackle head on what I take to be the fundamental weakness in full-blown cognitivist accounts, and it is necessary to do so by examining the difficulty in trying to understand emotions as beliefs - as ultimately understandable as an attitude towards a proposition.

Emotions and Propositions

If we accept the view that emotions can be understood as a collection of beliefs, and that they are therefore attitudes towards propositions, then we ought to be able to show that they can be analysed. This means that when we come to statements which employ emotion concepts such as;

(1) John is angry with James

we can analyse them into sentences of the form 'A believes that p'. The analysis can be made on the grounds that each emotion has a formal object, and this determines the description under which the object of any emotion must fall. As such, the analysis will involve replacing the emotion concept with a belief whose propositional content is determined by the formal properties of any object which falls under that emotion concept. Thus, we might argue that the formal object of anger is 'being wronged', so the analysis of (1) might read;

(2) John believes that he has been wronged by James

The first problem is that (1) could not be equivalent to (2), because we can be angry with someone even when we do not believe they have wronged us - it might just be that we are having a bad day, and our anger is directed towards the first person we happen to see. Here the cognitivist has a response based on Kenny's claim that the agent must hold certain relevant, true beliefs in order for it to be an *appropriate* emotion, so that it is an inappropriate case of anger if we falsely believe someone has wronged us or if we don't even hold the belief they have done something wrong. As such, the cognitivist can refine his claim to the effect that (2) is an analysis of John's being appropriately angry with James.

The more difficult problem, one which I see as insurmountable, is that even narrowing the claim to cover only appropriate instances of anger does not take us to the point where (1) is equivalent to (2), which it would do if it were a full analysis. We ought to be able to conclude that John's correctly believing that James has wronged him is equivalent to his being angry with James, but this just doesn't seem to follow. We can believe that we have been wronged, but the reaction simply needn't be one of anger. We can imagine scenarios in which we might react with indifference, regret, sadness, even amusement, but without anger. And it seems the same sort of problem potentially arises for any claim to equivalence between type (2) statements which set out the beliefs involved in the emotion concepts, and type (1) statements in which the emotion concepts are present. We can believe something is dangerous without fearing it, believe someone is worthy of love without loving them, believe someone is worthy of hate without hating them. At each turn, we can construct some sort of plausible account of why we can believe the type (2) proposition without experiencing the emotion in the type (1) statement. In other words, although emotions involve certain relevant beliefs, a full statement of the relevant beliefs need not involve the experiencing of the emotion.

If this is correct, then type (2) statements capture an important advance on the part of the cognitivist theories, in that we can conclude that it is a necessary condition for an emotion to be held appropriately that the agent has the true beliefs relevant to such emotions. As such, being afraid appropriately would entail that the agent hold certain beliefs about the fearfulness of a particular object or possible state of affairs. But the hurdle which the cognitivist still faces is that something is still lacking in the account, which is why we cannot conclude that holding the true relevant beliefs about the fearfulness of the object necessarily involves having the emotion.

A further technical problem emerges in the form of emotion statements which cannot be reworked into the form 'A believes that p'. The cognitivist literature is littered with examples in which statements of the 'John is angry with James' type are analysed, but there is a question as to whether or not all emotion statements can be recast in this way. Numerous emotions are described in sentences such as 'John hates James', or 'Mary loves French cooking', and there just seems to be something wrong in describing these emotions as attitudes towards propositions. Surely the object of the love and hate in

such statements are not propositions, but the physical objects referred to by means of them. If there is a discontinuity between the 'propositional' emotions and the rest, then we need an explanation of this, as well as a defence of the claim that all statements expressing emotions can be rewritten in the manner required for the analysis to begin.

There is a further problem which also confronts any cognitivist account, concerning the character of beliefs. The problem extends beyond the debate about the emotions, but it has particular relevance in this context due to the role of feelings in the debate. I have already said that I believe feelings to be partly constitutive of emotions, but I also intend to argue that feelings can be a part of the beliefs which may figure in an emotion, rather than a distinct phenomenon, and it is for this reason that no full analysis in terms of beliefs can be accomplished. I shall be looking at this in far greater depth in the next two chapters, but I wish to raise it first here in order to press home the shortcomings of cognitivism.

The kind of distinction I wish to draw can best be brought out by means of an example. Consider a young boy who holds the belief 'cricket is a great game'. He has acquired this belief from his father, who has instilled in him from a very early age the importance, tradition and beauty of the game, and the boy has always accepted this unquestioningly, if unenthusiastically. There then comes the point at which he plays in his first competitive match, and shortly after arriving at the crease to bat he receives a poor delivery to which he steps forward to play a flowing straight drive. He feels the impact of the ball on the centre of the bat, senses instantly that he has, for the first time, struck the ball with perfect timing, and sees it sail straight back over the bowler's head for a gigantic six. He experiences a surge of excitement, and the thought comes to him 'cricket is a great game', but this time accompanied by an intensity which was previously absent.

The question arises as to how we capture the difference which has arisen. Given that we are dealing with two cases of believing, then the propositional attitude remains constant. It therefore seems more plausible to describe the difference as one which can be seen in the content of the beliefs, and we might best describe the latter belief as one which is held *with feeling*, in a way which the earlier one was not. Much more will be said about this in the next chapter, but there are two points which I wish to highlight at this stage. The first is that there is an important and generalisable distinction between the beliefs

held before and after the ball flew over the boundary, and the second is that the distinction cannot be captured in terms of the semantics of the proposition. It is my contention that beliefs held with feeling partly constitute many of our emotions, and if the young cricketer were to express his joy after the match, this type of belief would be part of the emotion in this instance. If the nature of the belief has changed, and this change comes despite the propositional content remaining constant, then we must look beyond this for an understanding of the nature of feelingful and feelingless beliefs, and therefore for understanding of what an emotion containing such a belief is. To return to the cricket scenario, many emotions involve the kinds of beliefs held after the boy has hit the six and not before.

This criticism is more far-reaching than the general argument set out earlier that it always seems possible to hold the relevant set of beliefs and desires without experiencing the emotion. It means that it is not only a necessary condition to have feelings present for us to have such an emotion, but they must also colour the belief itself, rather than stand separately. This also works against Taylor's view that what we need is a better understanding of how the belief relates to the agent's wider experience and history which will make clear why some of us experience emotions in particular circumstances, whilst others don't. Once again, this may well be true on occasion, but it still won't do the required work to restore a fully cognitivist theory. In the story of the snake, what other relevant beliefs could there be beyond those which I have given. Both agents already have all the necessary true beliefs about the properties of the snake, and about the appropriate course of action. Unearthing more information about their past experiences with snakes may indeed help to explain why one of them is afraid and one isn't, but the explanatory force will not come in the form of additional beliefs, but in the form of how those beliefs are held. Surely we would want to provide an account along the lines of the fearless agent never developing any feelings of fear of snakes because he never learned to be afraid of them, or else gradually overcoming such feelings as his confidence grows in the company of others experienced in this way, and finally reaching the point where the fear he used to have exists only as a distant trace. The explanatory force comes not from holding different beliefs, but from holding the beliefs *in a different way*, and that difference is captured by the presence or absence of certain feelings.

One further point which applies to the kind of account offered by Taylor is that in the midst of a cognitivist theory, she nevertheless refers to 'what it would be human and natural for a person to feel under certain circumstances'.[26] Although I have criticised Taylor's overall account of emotions, I have also suggested that her emphasis on understanding them within a wider life is a crucial observation, but this is so partly because she offers what Michael Stocker calls a 'feeling-laden' account.[27] Her extended discussion of humiliation is at its most convincing precisely when she is drawing on material from James Joyce, in the same way as I used the extract from Lampedusa. But it is immensely difficult to conceive of such passages as an intricate series of beliefs – 'A shameful consciousness of his own person assailed him. He saw himself as a ludicrous figure, acting as a penny-boy for his aunts, a nervous well-meaning sentimentalist, orating to vulgarians,...the pitiable, fatuous fellow he had caught a glimpse of in the mirror.'[28] There is something within such writing which appears to capture the sense of humiliation which both Joyce and Taylor wish to bring out, but is this not precisely because it describes and explores certain feelings, rather than providing an adequate list of Gabriel's relevant beliefs? A similar point can be raised with regard to Solomon's writing, in that despite his explicit aim of providing an account which explains emotions without relying on the concept of feeling he nevertheless writes in such a way which often approaches a literary, feeling-laden style. His discussion of love in *The Passions* includes the following passage;

> But how difficult it is to be so vulnerable and trusting, to reject those many temptations to think of ourselves as superior rather than merely equal, to give up the successful defences and strategies that have worked so well for us in the past. But who can say of him/herself, "I don't want to love" or "I can't love", without the most profound regret. (277)

At an intuitive level, there just seems to be something wrong in trying to increase our understanding of sentences such as 'Concetta loves Tancredi' by subjecting it to conceptual analysis. The language of Joyce and Lampedusa does not simply describe, it evokes, and Stocker's concept of 'feeling-laden' description seems once again to be apt here. Insofar as our response to literature is unfeeling, then we cannot understand

---

[26] Gabrielle Taylor, *Pride, Shame, and Guilt*, 14.
[27] Michael Stocker, *Valuing Emotions*, 38-51
[28] Quoted from Gabrielle Taylor, *Pride, Shame, and Guilt*, 9.

Gabriel's humiliation or Concetta's despair. We would simply have a flat description of something like mental states which would not engage us in a way which we would want to call an understanding of the texts. Our distance would cut us off from the intended meaning. If this holds true for literature, then it tells us something much wider about the emotions in general. The attempted analysis of emotions in terms of beliefs and attitudes or desires leaves out a key concept and the means for understanding an emotion. The way in which we speak of the emotions is itself important, and this is because the feelings which the cognitivists leave out in analysis are precisely what is required for us to experience an emotion, as well as understand what it is for someone else to experience one. Once we are restricted to beliefs, in whatever combination, one cannot successfully express or understand what an emotion is.

Conclusion

It is now time to focus more specifically on the concept of feeling, and the role it plays in our emotions, which will be the subject of the next chapter. But it is worth highlighting the aspects of the cognitivist theories which must play a key part in our understanding. The idea of emotions having objects is one which is accepted even by those who would reject the full cognitivist account. Our fear, anger or love is always for someone or something. This property of intentionality is integral in the role emotions play in our lives, in that it is through the emotions that the outside world takes on the significance that it has for us. The parts of the world which we fear and the people whom we love are not contingent aspects of our survival, but central features of a life, and the specific objects of our emotions therefore take on immense importance in light of this.

The causal role of beliefs in our emotions are also central, in that they systematically provoke and alter emotions, often in immediate and dramatic fashion. If I am informed that my beloved car which is missing has been stolen by the friend who is now comforting me, feelings of despair and gratitude can be instantly transformed into righteous anger. If my emotions remain unaltered in light of new beliefs, as in Kenny's example of my envy of what turns out to be my own fruit trees, then the emotion can be deemed inappropriate. My beliefs can also influence the appropriate intensity of my emotions, in that the degree of my fear ought to alter depending on whether I believe I have lost one hundred or one million pounds on an ill-judged stock market deal. As

such it is important to understand emotions as complex phenomena, and to resist the tendency to reduction. This complexity is not only in terms of the various elements which go to make up an emotion, but also in the importance of how many of our emotions must be seen as part of a wider life, and fully comprehensible only in that context.

Once we have established the crucial part of beliefs in our emotions, then it remains to bring them into the right juxtaposition with the related feelings, and it is to this which I now turn.

2. Feeling

Introduction

The aim of the last chapter was to highlight weaknesses in accounts of emotion which try to dispense with the concept of feeling. The tendency amongst cognitivist thinkers is to see feelings as a distraction to a proper analysis of the emotions. Kenny, for example tells us that 'feelings of emotion are the sensations linked with the symptoms of an emotion;'[1] The aim of this chapter will be to argue that feelings are absolutely central to the experience of having an emotion, and to the understanding what an emotion is.

Discussion of feeling is complicated by the different senses in which this concept is used, and it is as well to identify two which are *not* relevant to my usage here. The first is feelings related to our sense of touch in statements such as 'I felt the roughness of the old fabric'.[2] The second is that of 'proprioception', meaning our awareness of our own body and limbs, and their spatial location with regard to one another, as in statements such as 'I felt my arms hanging limply by my side'.[3] The concept of feeling to which I shall refer when defending the claim that feelings partly constitute an emotion is used in phrases such as 'I feel tired' or 'a feeling of satisfaction'. I shall treat such feelings as perceptions of the state of one's own body,.[4] and unless otherwise stated, I shall use both verb and noun in this sense. One further important distinction which needs to be made is between bodily feelings and bodily sensations. I shall take bodily sensations to be characterised by their locatedness in a specific part of the body, as in cases of one's eyes stinging, one's back aching or one's temples throbbing. Bodily feelings typically indicate a perception of the physical state of the body as a whole, as in cases such as feeling fatigued, where it is the whole body which is the object of the perception. Feelings thus tend to be perceptions of more generalised states of the body, but they may well include sensations, such as in cases where I feel hungry, and this involves the sensation of my stomach rumbling.

---

[1] Anthony Kenny, *Action, Emotion and Will*, 99

[2] For a discussion of this topic, see M G F Martin, 'Sight and Touch', in *The Content of Experience*, ed T M Crane.

[3] For discussion of this concept, see Brian O'Shaughnessy, 'Proprioception and the Body Image', in *The Body and the Self*, eds Bermudez, Marcel and Eilan.

[4] In treating feelings in this way, I follow the line in recent treatments of this concept in philosophy of mind. See for example D M Armstrong, *A Materialist Theory of Mind*, Chapter 14; M G F Martin,

In order to shed light on the concept of feeling, I shall look at several issues concerning the nature of feelings which partly constitute emotions. Firstly, I shall be looking at the distinction which has been drawn between 'bodily' feelings on the one hand, and 'psychic' feelings on the other, which broadly equates to the contrast between feelings which pick out physical states on the one hand, such as feeling physically drained, and on the other hand the feelings associated with boredom (where we are aware of no sensations or bodily feelings). Next, I shall consider the relationship between experiencing feelings and our level of awareness of them. We see this sort of possible distinction in cases where we might talk of our fear of a bully, something which we think of as an ongoing state, and the kind of fear we might experience when we see him approaching us, when the kind of bodily feelings described above may well afflict us. We need to be clearer about the kind of feelings involved in both sorts of experience, and about the different levels of awareness of these feelings which we will have.

A further question arises as to whether or not feelings must be present at all times during an emotion, or whether or not it is sufficient for them to be present on certain occasions in order to describe the experience as an emotion. This is important in that we often speak in terms of emotions which last over extended periods of time, such as envying someone for years. If feelings partly constitute emotions, then must we be committed to the view that we have experienced certain feelings over a period of years in order to justify the claim we have envied someone all this time? Finally, I shall consider the concept of 'feeling towards', and the idea that our feelings have an object, as against being purely internal phenomena.

Broadly speaking, the position which I shall be arguing for on each of these questions is that we require a very broad concept of what feelings are which can take account of their multifarious nature. We need a conceptual scheme within which we can accommodate the feelings associated with extreme fear, which include bodily sensations such as the thumping of one's heart, as well as the more subtle psychological state of feeling pride at a clever chess move. We need to recognise that feelings can wax and wane over time, and that we will not always be fully aware of what we are feeling,

---

'Bodily Awareness: A Sense of Ownership', in Bermudez, Eilan and Marcel, edd. *The Body and the Self*; Brian O'Shaughnessy, *The Will*, Vol 1, Chapter Five.

meaning that we may be feeling something without 'attending to it'. There are also experiences where we feel nothing in circumstances where this would not normally be the case, as with athletes who claim to have felt no tiredness as they were striving to cross the finishing line in an important race. I shall also be trying to show that the key to understanding how our emotions bind us to the world around us is the concept of 'feeling towards'. It is our feelings towards persons, objects and activities in the external world which force us into an engagement with the world which we find significant. This will be the conclusion which I hope emerges from the arguments which I shall offer on each of the issues listed above, but I shall also introduce one of the key concepts in outlining the relationship between our emotional life and questions of value. I shall argue for the existence of what I have called a 'background of directed feeling', which is an acquired reactive ability to respond to the world in one way rather than another. This will form the foundation of the later theory, but first requires a fuller exposition of the more basic concept of feeling.

Bodily and Psychic Feelings

Of the two categories of feeling under discussion here, it is clear which is the easier to identify. The bodily feelings and corresponding sensations normally associated with emotion are those such as tightening of the muscles, increased heart rate, sweating, trembling and other changes to the autonomic nervous system which we experience as a part of a wide range of different emotions. An initial point which is required for the sake of clarity is that feelings do not correspond to bodily changes, but only to those bodily changes which can be *felt*. As such, they don't include the increased speed at which the blood sometimes races through my body, but they do include the perception of the increased speed at which my heart is beating, which is causing this. It is also worth reiterating the point from the last chapter that we cannot say that any particular bodily feelings indicate the presence of a particular emotion. Indeed, if we consider extreme examples of any emotion then it is usually possible to think of scenarios in which the more common bodily feelings might be present. I may feel extremely tense at moments such as when William James' elephant comes racing towards me, or on discovering that my car has been stolen, or on seeing the woman I have worshipped for years at the theatre with my long-time rival for her affections. It may also be impossible to attribute such feelings to a single emotion. If we take the last example of my seeing the couple at

the theatre, my emotional state may be a combination of intense jealousy at the sight of the two together, as well as extreme anger at his success.

A further characteristic of bodily feelings is that they tend to provide a good indication of the intensity of an emotion. Although specific bodily feelings may not be correlated to specific emotions, their appearance tends to justify the claim that the emotions are more intense. If we take the example of fear, then feelings which include the sensations of my hands trembling and my accelerated heart rate just couldn't reflect a mild fear of rain disrupting a casual tennis match. Such changes only make sense in cases where a standard agent (where 'standard' excludes exceptional cases such as phobics) perceives the object of the fear as having a potentially important impact, perhaps due great danger, or damage to a vital project.

It would be a mistake, however, to assume that all extreme emotions must have some sort of bodily feeling as a constituent part. One can easily imagine someone full of pride at a particular achievement, without assuming that her heart is racing or her knees shaking. Similarly, one might spend several hours being angry with someone, with the anger erupting from time to time when one thinks back to the original incident, but it would be unusual for one's heart to be racing for what we take to be the entire duration of the emotion. So, to describe an emotion as intense will be justified in terms of the strength of the feelings of the person who is having the emotion, but there is more which can be said here. We can explain why the feelings (and therefore the emotion) are more intense by considering the other constituent parts of the emotion. Great fear will be partly characterised by the bodily feelings, and partly explained by the beliefs of the agent with regard to her own predicament. It is when we believe we are in great danger, or that some highly valued aspect of our life is imperilled that the fear will be intensified. This means that whilst the intensity of our emotions may have powerful bodily feelings as their most emblematic feature, our system of beliefs and values are also central in understanding intensity in emotion because they provide the contextual framework within which we perceive what is to count as very important. If someone is panic-stricken at the thought of a fall in share prices, then this heightened emotional state reflects a way of perceiving the world, and the intensity of the emotion only makes sense in light of a highly personalised range of beliefs and desires related to this eventuality. Intensity must therefore be understood as a feature of emotion which is grounded in,

and reveals, the place which the object of the emotion has in the agent's scheme of beliefs, desires and values. In some cases, this will reveal very little, in that many of us will share a similar range of beliefs, desires and values with regard to remaining alive, and showing intense fear in the face of great danger will therefore reveal very little we didn't already know. But in other cases far more will be revealed. Great fear with regard to share prices falling, a particular football team facing the prospect of relegation, or a sudden upsurge in the popularity of utilitarianism could only possible be understood in the context of the agent's particular engagement with the world.

The concept of bodily feeling stands in marked contrast to that of 'psychic' feeling, which is of crucial importance in the work of Michael Stocker.[5] Stocker follows Descartes in describing psychic feelings as those which 'we feel are in the soul itself', as against bodily feelings which we perceive as a condition or change in the body. A bodily feeling is the perception of an overtly physical state such as being physically tense, whereas psychic feelings do not have this physical character. An example of the contrast comes in comparing the bodily feelings of rage with the psychic feelings of pride. The points at which psychic feelings are perhaps best highlighted is in the phenomenology of areas of our lives such as interest, care or concern. These are aspects of our mental life which can be more or less constant over long periods, with the same person or activity for their object, and without necessarily including the bodily feelings which are associated with emotions such as rage or terror. Examples of where such feelings are present could be seen in cases such as a novice priest who is determined to make it through his training in order to practice his vocation, or the teacher who has an ongoing pursuit of bringing the best out of her students. The reason such examples are of importance is that they hint at the existence of subtle, stable and guiding feelings which contrast with those more prominent ones in our violent emotions. They provide a background motivational force which can drive people forward in the pursuit of long-term projects and commitments in a way in which the extreme, short-term feelings of rage cannot.

If we wish to characterise more precisely what psychic feelings are, then I suggest that they can best be understood as perceptions of a certain class of non cognitive mental states. With bodily feelings, we perceive overtly physical states such as the tension of our

---

[5] See Michael Stocker, 'Psychic Feelings: Their Importance and Irreducibility', in *Australasian Journal of Philosophy* 61, 5-26, and *Valuing Emotions*, 17-54.

muscles, but with psychic feelings we perceive mental states such as satisfaction, unease or displeasure, those very state which we standardly describe as feelings, but which do not apply to the kind of physical states which fall under the heading of bodily feelings. An example will help to clarify this. Consider a scenario in which you have been working on a particularly thorny problem in symbolic logic for several hours, and you find you are starting to make sloppy mistakes which are costing you considerable time and effort. One way of understanding the problems you are now having is to think that you feel tired, but this sort of mental tiredness contrasts sharply with the physical tiredness one feels after a running a marathon. The perception is of a diminution in one's *mental* capacities, including features such as an inability to concentrate and an increase in the level of determination and mental effort required to perform to one's normal standard. In other words, we become aware of what we would tend to think of in folk psychological terms as a mental rather than a physical state, and this perception can be characterised as a psychic feeling.[6] With this view of psychic feelings we are better placed to understand emotions such as pride, where there are not necessarily any bodily feelings, but rather psychic ones which reveal the way our own beliefs seem to us.

If we retain the example of pride at a particularly successful chess move, then the belief about the quality of the move and the identity of who played it is combined with a feeling of satisfaction which colours the beliefs. This is not the bored reflection of a grandmaster who has played a thousand such moves before, but the impression of someone for whom the achievement resonates in a particular way. It is not simply that someone has played a great move, but that *I* have played a *great* move. The satisfaction influences our perception of ourselves in light of the quality of the move. Similarly, if I am angry with a noisy neighbour, the beliefs about the source of the noise and the distraction it is causing is coloured by the displeasure which is combined in my consciousness with the beliefs.[7] In this way, emotions can be the result of the combined efforts of my beliefs and feelings, where the feeling is the perception of a mental state such as satisfaction or displeasure. This means that the overall account of the nature of emotion can be retained, but now with a more fine-grained account of the kind of

---

[6] It is worth emphasising once more that this needn't indicate a dualist position, in that the option is still open for the materialist to argue that this phenomenon will be fully explicable in terms of the operation of the brain.

[7] I shall have much more to say about the way in which beliefs and feelings are combined in consciousness in the final section of this chapter.

feelings which are involved in emotions. The feelings which are combined in consciousness with beliefs of perceptions may be bodily or psychic ones, or both.

Even with a clearer idea what a psychic feeling is, we might still question whether or not it need play any sort of necessary conceptual role which could not be filled with a desire. With the case of both the priest and the teacher, why couldn't we simply attribute certain beliefs about the priesthood and teaching, and certain desires with respect to potential achievements. There just doesn't seem to be any *prima facie* need to attribute any feelings, except in those moments where there is a distinct emotional experience, including a particular bodily feeling. This sort of response becomes more attractive given that one could argue that we would often appear to be unaware of these feelings if, as I have suggested they might be present over a period of years, motivating us without our recognising them on each occasion. The danger is that they start to look like rather mysterious phenomena, hidden from view in the unconscious, and invoked purely to plug a perceived conceptual gap.

One response would be to treat this as an attempt to reduce psychic feelings to desires. This cannot work if one accepts the characterisation of feelings as perceptions of a bodily state, in that 'I feel that I am tired' cannot be analysed into 'I desire that I am tired'. Indeed, if we try to construe the contrast between beliefs and desires along the lines of 'direction of fit', feelings are going to come out closer to beliefs than desires, in that both can be understood as aiming at fitting the world, rather than aiming at having the world fit them.[8] But this doesn't get us off the hook entirely. What if instead of saying we can reduce psychic feelings to desires, we argue that there is nothing there to reduce? Stocker's discussion deliberately provides psychic feelings with a motivational force in their role of caring and showing concern for persons and projects over the course of our lives, and couldn't this role could simply be played by a desire? We could concede that loving someone over many years does means having some bodily feelings but, but then argue that *acting* out of love or concern requires the desires that such feelings might provoke, and it is these desires that are doing the ongoing motivational work.

---

[8] For a fuller discussion of this, including qualifications on how useful this schema is, see Peter Goldie, *The Emotions*, 24-26.

By way of response, one can begin by offering a straightforward formal argument for the existence of psychic feelings. If all emotions are partly constituted by feelings, and not all emotions involve bodily feelings, then some other sort of feeling must be present in some emotions. Affective states such as care or concern, as well as long-term emotions such as an enduring love could be said to involve subtler feelings such as contentment, satisfaction, refined pleasure, and these are the sorts of states which should be understood as psychic feelings.

The next move might come in the form of a limitation in the explanatory potential of using only desires in a theory motivation, as against using the concept of feeling in addition. There is a distinction between 'cool' desires, and 'hot' desires, or what I shall call desiring with feeling which can be brought out in a parallel example to the young cricketer used in the previous chapter. Imagine that prior to hearing any classical music in your life, you have been told by a friend that you would enjoy attending a performance of *Madame Butterfly*. You accept this advice in a very coolheaded manner. A desire then forms to attend a future performance, and in order to draw maximum benefit from the experience you buy a CD and listen to the opera for the first time. You are overwhelmed by the power of what you hear, and the desire to see the opera comes to you once more. Just as with the cricketer's belief about the value of cricket on experiencing the high of succeeding at it, your desire to see the opera may be qualitatively different after hearing the CD, and the best way of describing the difference to focus on the contrast between feelingless and feelingful desire. If we acknowledge that it is possible to hold the kind of desire with feeling that I have just described, without necessarily having any bodily feelings, then we have a further argument for the existence of psychic feelings. Not only this, but the kind of feelingful desire described above may well have a greater motivational force than its feelingless predecessor. We might now find the desire to attend the opera has greater force and urgency, and that it now dominates other desires and preoccupies us to a greater degree.

If these arguments go through, then we are entitled to conclude that there are such phenomena as psychic feelings, they are distinct from desires, they can come to influence desires, and they can do so in a way which adds greater motivational force. This is not to say that stronger feelings *necessarily* add to the strength of our motivation to act. One could be glowing with pride, with the most intense feelings of satisfaction involved in the

experience, and yet have no desire to act at all. Conversely, one might have quite mild feelings associated with fear, and yet a strong desire to leave the area. These conclusions are preliminary ones, designed to set the scene of the claim that we have 'a background of directed feeling' which grounds our desires, and which helps to explain why we have the desires we have. This will be the focus of the next chapter, but I now wish expand the account of what feelings are by considering in greater detail the role of feelings within an account of how emotions succeed in motivating us.

It is, I think, uncontroversial to claim that emotions can motivate us to act, but we now need to get clearer about the role of belief, desire and feeling within the motivational story we are going to tell about emotions. When we have statements such as 'I proposed because I love her', 'I hit him because I was angry with him', or 'I ran because I was scared', we see examples of actions which one might say are actions 'out of' love, anger and fear. But what exactly do we mean when we say we acted 'out of' the emotion? In these sorts of descriptions, I suggest we describe an action, and in referring to the emotion in this way we are rationalising that action, by which I mean we are stating the reason for it.[10] Support for this sort of view can be given by observing that if we ask an agent why she performed a particular action and she offers the kind of answer set out above, we take this to explain why she acted as she did - she has given a reason which allows us to make sense of the action.

Now, how can this sort of Davidsonian account of reasons and actions be applied to this theory of the emotions? Davidson argues that having a reason is a combination of having a belief and a pro attitude, with a pro attitude comprising a wide range of attitudes such as 'desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values in so far as these can be interpreted as attitudes of an agent directed towards actions of a certain kind'.[11] In this sort of scheme, we might say that emotions are examples of pro attitudes, and this may well be what Davidson thought, as he included

---

[10] I use the same formulation as that found in Peter Goldie, *The Emotions*, 37-47.

[11] This is the term used by Donald Davidson in 'Action, Reasons and Causes', *The Journal of Philosophy*, 1963, 685-700.

[11] *Ibid*, 686.

'love of children' as one example of what a pro attitude might be.[12] But this would sit uncomfortably with the view that emotions are partly constituted by beliefs, in that it looks as if the same beliefs may come up twice in the explanation of a particular action. In a statements such as 'I proposed because I love her', the account of the emotions as complex phenomena partly constituted by beliefs would mean that the love might involve beliefs such as 'she is the kind of person I would like to marry', or 'she is just the sort with whom any man in his right mind would like to spend the rest of his life'. If we then move to our explanation of why we acted, and we take the emotion to be the pro attitude, we must find the beliefs which combine with the pro attitude to provide the reason for proposing. These beliefs would surely be just those which we have already identified as part of the emotion.

In light of this, the most plausible account is one in which we see emotions as phenomena which involve pro (or con) attitudes. Fear involves a con attitude to the object of our fear, love a pro attitude to the object of our love, joy a pro attitude to the object of our joy, and so on. This much seems clear from the ways in which we might describe what such emotions involve when we speak of them, such as when we say 'my love involves a longing to be with her' or 'my fear of heights involves a dislike of going anywhere near cliffs and other high places'.

The question now becomes how pro attitudes relate to beliefs and feelings within the emotion. The answer to this emerges when we take account of certain properties of feelings - every feeling has the property of being either pleasant or unpleasant,[13] which is surely a sufficient condition to generate the pro or con attitude which is also a part of the emotion. Once again, this receives support from some common sense observations about such emotions, such as when we talk of the torment of jealousy or the wonder of love. Given the premises that all emotions involve feelings, all feelings are sufficient to produce pro or con attitudes, and that all such attitudes are capable of prompting action, then one can argue that all emotions are reasons for action. Like Davidson, I take reason

---

[12] *Ibid.*

[13] Even though we speak of 'feelings of indifference', it is misleading think of these as a neutral feelings. It may mean we have no feelings about something or someone, in which case it is not that we have a neutral feeling, but that there is no feeling to be pleasant or unpleasant. The alternative is that when we say we are indifferent to someone it is a euphemism for saying we see them as being, so to speak, below the radar which picks up those who count - it is an insult.

giving explanations as a species of causal explanation, and we can therefore add that emotions cause the actions for which they are reasons.

The problem with a claim of this force is that there is an obvious counter claim that we may well experience an emotion, but not act out of it, and even claim that we had no reason to act. If I recognise in myself an overwhelming desire to acquire as much money as I can, and I know that marrying the woman I love will result in my being written out of my rich grandfather's will, then I may decide that love is not a reason for me to act (in the form of marrying her) in these circumstances. But to say that I have not acted out of love does not mean that the love had no motivational force, or that it could not have been a reason for action under other circumstances. On the contrary, once someone states that she is experiencing an emotion, then we generally assume that this is a fairly standard motive to act, and we would accept this as an explanatory reason for an ensuing, relevant action. In cases where we don't act out of the emotion, this must be because there are other reasons which have outweighed the motivational force of the emotion in this instance. In light of this, we simply need to qualify the stronger claim that all emotions are reasons for action, such that the claim becomes; all emotions are reasons for action, but reasons which may be outweighed by other reasons where there are overriding desires or other pro attitudes which have stronger pull than the emotion in this instance.[14] Davidson makes a similar point with regard to the motivational force of desires;

> Any serious theory for predicting action on the basis of reasons must find a way of evaluation of the relative force of desires and beliefs in the matrix of decision; it cannot take as its starting point the refinement of what is to be expected from a single desire.[15]

---

[14] This commits me to what I take to be the standard Humean line that all desires (meaning pro attitudes in this context) can be reasons for action, and can only fail to motivate us if there are stronger reasons pulling us in a different direction.

[15] Donald Davidson, 'Action, Reasons, and Causes', , *The Journal of Philosophy*, 1963, 697. It is perhaps worth pointing out that one could accept this explanation of how emotions motivate us without accepting the general belief desire account of motivation. One might wish to argue it is a feature of emotions that they generate desires (where this is taken in the very general sense of a pro attitude), and that emotions motivate as a result of the desires they give rise to, but this does not indicate that actions not out of emotion therefore require desires as part of the explanation of why the agent acted.

There is a further point worth raising at this stage. We need to be clear about the distinction between acting 'out of' emotion,[16] which I am discussing here, and acting 'for the sake of' an emotion. Acting out of love will mean that I have a series of dispositions, feelings, concerns, thoughts etc which mean I am directly concerned with the object of my love. To act for the sake of love will mean that I am directly concerned not for the object of my love, but for the love itself. Where this is important is that one might act for the sake of love, such as in cases where the love has gone dead and we are trying to revive it, but not have the dispositions, concerns, thoughts etc characteristic of those who are acting out of love. This means that although there may be descriptions of specific actions which could be similar when it comes to acting for the sake of an emotion and out of emotion, there will be points of fundamental divergence in the phenomenology of the agent. Many of the thoughts, feelings and intentions of the agents will differ markedly, and this is precisely because they are not experiencing the same phenomenon. This will be of crucial importance later on, in that I shall be claiming that the virtuous agent is one who acts out of compassion. That is to say she is someone who has a direct concern with regard to addressing the needs of others as revealed through their suffering, and acts out of that concern. When asked why she acted as she did, she will give as her reason that it was the compassionate thing to do, and this will rationalise the action. This will ultimately provide us with an account of how it is we come to be motivated within the moral domain, but it is now necessary to fill out in greater detail the description of the nature of feeling by considering to what extent we are aware of these perceptions.

The Awareness of Feeling

I have argued that feeling is a perception of the state of one's body, but there is a series of questions which need to be answered with regard to the relationship between experiencing a feeling, and being aware of that experience. Consider for example the case of a soldier who has been shot in the arm, but continues to fight and later says that he was unaware of the wound at the time, or of an athlete who struggles for the finishing line and claims not to have felt any exhaustion despite the intense physical demands of the race. Let us say that we observed certain changes at times when they claimed not to have the feelings we normally associate with such behaviour, such as the soldier using the

---

[16] For a full discussion of this discussion, see Michael Smith, 'The Possibility of the Philosophy of Action',

arm in which he has been shot less than before, or the athlete slowing down. In either case, is he feeling something, and if so, is he aware of this feeling?

The line which I shall take is that we need a conception of feeling which is as broad as that other perceptual faculties such as sight or hearing. There are cases where we would normally expect to see something, but fail to because we are distracted, or else we sometimes see things without 'attending to' them. I shall try to show that the same is true of feeling. In addition, awareness of what we feel needs to be thought of as a state which admits of degrees, so that we can be aware of an object without being what one might want to call 'fully aware' or 'fully conscious' of it. In each case, behavioural changes will help us to attribute feeling or awareness from the third person perspective, even where the first person perspective indicates the subject is not fully aware of what she is feeling.

As a starting point, lets us pursue the analogy with visual perception. We sometimes travel home by car and realise we have arrived without having concentrated on the journey, usually because it is such a familiar route. Nevertheless, we would acknowledge that we certainly saw the red light at which we stopped, and then the green light which prompted us to accelerate away, and we could justify this claim with reference to our behaviour. If one day we failed to slow down at the red light and crashed into the car in front, then we might plausibly explain this with reference to certain mental states which prevented us from seeing the light, such as our preoccupation whilst driving with a work-related issue. In other words, our mental focus can prevent perception in cases where this would normally occur.

The analogy with feelings now starts to become clear. If we take the case of the soldier, his level of concentration on slaughtering the enemy in front of him has prevented him from feeling the injury sustained as a result of the bullet wound. Once the distraction is removed, he will begin to feel pain as a result of the injury – to perceive the state of his body. But what of the observation that the soldier had started to protect the wounded arm, or that the athlete had slowed down? Don't such observations suggest that they must have been feeling the pain, or else they would have continued on as before? I don't believe this follows. I drew a distinction at the beginning of this chapter between bodily changes and bodily feelings. Although feelings will be perceptions of changes in the

---

reprinted in Bransen and Cuypers eds. *Human Action, Deliberation and Causation*, 23-26.

body, not *all* changes in the body will be felt. In cases such as those under consideration here, the body may respond without our deciding it, just as in cases where an illness results in lack of hunger. We may not feel constantly sated, or even be aware of the fact that we haven't eaten for twenty-four hours, yet the behaviour could be identical with that of someone who does feel sated, and is fully aware of this. Feelings cannot simply be read off from behaviour, but need to be attributed after considering behaviour as but one criterion for signalling their possible presence. If this is right, it also means that feeling is a conscious state, which brings with it a conceptual requirement that there be at least some level of awareness of it, even if we are unable to recall it afterwards. On this view, to say that we could not be in pain and be unaware of it is simply one example of not being able to feel something without being aware of it.

This in turn requires us to consider a feeling as one of many perceptions we may be having at the same time, all of which will command our attention, but to greater or lesser degrees In the case of the driver who returns home 'on autopilot', we must surely attribute at least a low level of attention to his visual perceptions on the grounds that he must be able to interpret what it means for the lights to change in order to produce the appropriate behaviour To put it another way, he must not only see the lights turn, but also be aware of what this signifies in order to slow down and speed up. But this contrasts sharply with the thought which may well come to him as he hits the back of the car in front, looks up at the lights, and thinks to himself 'My God, the lights are red!', where the level of attention to the perception will be extremely high, as his concentration switches dramatically from work issues to the redness of the traffic light.

I believe it is appropriate to acknowledge a similarly broad range of levels of attention to our feelings. We can be vaguely annoyed at a dripping tap, or else find ourselves utterly dominated by feelings of irritation which we try to suppress but which continually fight their way to the forefront of our minds. In order to elucidate these last points, I shall use two extracts from Michael Ondaatje's *The English Patient*. The first extract describes a encounter between Kathrine Clifton and Almasy, who are shortly to embark upon a passionate and tragic affair. The second extract is one of Almasy's reflections about the relationship itself. Clifton is Kathrine's husband;

Extract 1

She had consumed all her reading and asked me for books. I had nothing but maps with me. "That book you look at in the evenings?" "Herodotus. Ahh. You want that?" "I don't presume. If it is private." "I have my notes within it. And cuttings. I need it with me." "It was forward of me, excuse me." "When I return I shall show it to you. It is unusual for me to travel without it."

All this occurred with much grace and courtesy. I explained it was more a commonplace book, and she bowed to that. I was able to leave without feeling in any way selfish. I acknowledged her graciousness. Clifton was not there. We were alone. I had been packing in my tent when she had approached me. I am a man who has turned my back on much of the social world, but sometimes I appreciate the delicacy of manner.[17]

Extract 2

She had been part of the expedition for almost a year. I saw her, conversed with her. We had each been continually in the presence of the other. Later, when we were aware of mutual desire, these previous moments flooded back into the heart, now suggestive, that nervous grip of an arm on a cliff, looks that had been missed or misinterpreted.[18]

I suggest the recognition of the feelings which had been present for some time prior to Almasy's acknowledgement of his intense desire for Kathrine is typical of the way in which feelings can erupt in such a way which highlights their forcing their way to the forefront of our minds. It is not that there were no feelings present before, but that they are now stronger and demand greater attention. Typical examples of how this occurs in ordinary life could be irritation at a neighbour with whom we have always felt constrained to be polite, or anger at having constantly missed out on a promotion which we believe should have been ours, or, as in the case of Almasy, feelings of love which seem at odds with the wider range of qualities which we perceive ourselves as having. Such feelings could be either bodily, or psychic, or both. In the case of Almasy, I suggest that the combination of intense romantic and sexual desire, coupled with the extraordinary break with the rest of his life, constitutes the kind of emotional experience which Goldie describes as one in which '[o]ur entire body is engaged in the emotional

---

[17] Michael Ondaajte, *The English Patient*, 231
[18] *ibid*, 235

experience, and all the feelings are "united in consciousness" in being directed towards its object: united "body and soul", "heart and mind."'[19]

Much of this suggests that when we talk of awareness of our feelings, this must be understood as potentially involving much more than simply registering their presence. Once feelings are acknowledged in the way described in the second extract, feelings which were held before can now be better understood. When Almasy says that previous moments 'flooded back into the heart, now suggestive', we can interpret him as now being able to explain moments such as the discomfort he felt during the exchange over Herodotus as the nascent feelings which would ultimately develop into the intense romance he was later to experience. This is not to say that he can look back and now say he was in love with her at that moment, which would be a much more controversial claim. But it would be possible to understand the feelings of discomfort as the onset of a shift in his perception of himself, caused by Kathrine, which challenged the detached view of himself he previously held, and signalled the move towards the intense desire which would come to characterise his attitude towards her later on.

If this is correct, then awareness of feelings can involve coming to an understanding of past feelings which we now see in a different light, or else a sudden and blinding recognition of feelings which were present before, but of which we are only fully aware now. A further possibility would be one in which the initial feelings between Almasy and Kathrine did not erupt into a passionate affair, but developed gradually over the course of many conversations or encounters during which a range of different shared interests and pursuits brought them a steadily growing awareness of increasing attraction. In such circumstances it would be misleading to limit oneself to the simple categories of feeling nothing, feeling something and being vaguely aware of it, and feeling something and being fully aware of it. There may well be many degrees of awareness through which we pass before we can say we are fully aware of feelings associated with love.

A more difficult question is how we would describe a situation in which the feelings were ones of which neither protagonist was ever fully aware. Let us say that they had both been killed after the scene described in the first extract. Could we then say that they were attracted to one another, or that they were in love? It is not clear to me that there

---

[19] Peter Goldie, *The Emotions*, 55

could be a confident answer to this question. The uncomfortable way in which Almasy addresses Kathrine, and the understated reflection about appreciating the delicacy of manner indicate his discomfort, but is this enough for us to attribute a full-blown emotion predicate such as 'in love' or even a more modest one such as 'attracted to'? I am inclined to say it is not, on the grounds that we require a greater degree of awareness of our feelings, and more evidence in terms of our behavioural responses before we can use such predicates. But I would also concede that other people's intuitions may run in the opposite direction, and it may be the case that feeling and emotion concepts just aren't fine-grained enough to deal with these sorts of borderline cases.

This last point brings us to the difficulty in describing feelings. There may be instances where I can specify a particular time at which my fear started, and this might be taken as the point at which my heart started beating on perceiving some imminent danger. But if we are discussing the more subtle cases of psychic feelings involved in more complex emotional states such as love then it is much more difficult to talk in such precise terms. Discussion is further hampered by the fact that we have such a limited vocabulary. We can pick out certain properties of many of our feelings such as pleasant or satisfying, but these provide little more than the barest of descriptions, and we are better off turning to literature to gain clearer insights here.

Nevertheless, we can certainly gain a much a clearer picture of the role feelings play in emotions, and one of the crucial aspects of this account, which is also important if the references to levels of awareness are to count, is that there is an increasing reliance on the concept of 'feeling towards'. If there is to be an analogy between decisions which I take without concentrating which guide my driving, and feelings I have without attending to them which guide my behaviour, then it must be shown that the feelings are directed towards the world in an analogous way to the one in which judgements are representations of the world. Although I have defended the idea of the intentionality of emotions, the intentionality of feelings is a distinct issue which requires support if the other arguments offered here are to stand up. This is of particular importance in the context of the theory as a whole, in that I hope to show that it is our feelings towards the world which ground our meaningful commitments and relationships, including our moral commitments. But if this is the case, then it will first be necessary to establish that feelings can have the stability and consistency which would allow them to play this role.

The Constancy of Feeling

One of the claims which I have defended throughout this chapter and the last, is that feelings are necessary constituents of emotions. I have argued that in cases such as where we believe someone is dangerous, and desire to escape him, this does not entail fear - a feeling must also be present. One problem which emerges from this is just how we are to deal with statements such as 'I have been in love with my wife for twenty years' or 'I had an argument with my brother last week and I have been angry with him ever since.' These are the kind of statements about emotions which we find entirely acceptable in ordinary usage, but how then should we accommodate the claim that emotions must have feelings? It seems we must either argue that the man who is in love with his wife has had a constant feeling for twenty years, or else the feelings have come and gone, and each option presents serious difficulties. To the man who claims to have had a constant feeling of love over the entire period, we might ask if he felt love while he was storming out of the house after a furious argument, or if he had feelings of love while he was asleep. To the person who has been angry for a week, who contends that the feelings involved have come and gone, we might ask whether this is one episode of anger, or whether or not he should say not that he had been angry for a week, but that he had been angry on the different occasions when the feelings had come to him.

I shall look at these alternative positions and reject both in favour of a position in which references to enduring emotions need be seen as entailing neither the claim that the feelings were permanently present, nor that we need to dispense with the idea of long-term emotions altogether. Instead, we need to think of enduring emotions as ones in which we have feelings which may wax and wane, and of which we may be more or less aware over time. There may even be points when we might think of the feelings as disappearing altogether, as in cases where we fall out of love with someone for a period, but we can still justify the claim that people experience emotions over many years by thinking of certain feelings as having been present, and having been sufficiently strong and enduring to characterise our attitudes towards a particular individual or event, and thereby allow us to speak of loving a person or being angry with someone over a period of years.

The idea that feelings cannot possibly be present at all times during an emotion is an obvious conclusion which follows from the fact that we refer to periods during which we claim to have emotions which extend well beyond the periods during which we are conscious of having feelings related to those emotions. To take the example of the anger which lasts a week, the typical phenomenology of such a state would be an intense period at the beginning where we might combine bodily and psychic feelings, and during which our thoughts may well be utterly dominated by our anger towards the person who has antagonised us. This may be followed by a period of hours when we still feel on edge, and we often return to the central issue and the words which were exchanged. After this, we may gradually, perhaps over a period of days, think of the incident less and less, and allow other topics to concern us. Yet even in this final, more sedate phase, we might well still describe ourselves as angry, despite the feelings being fully consciously present only intermittently and to a much weaker degree.

In light of this, we could claim either that the feelings are dispositional, or that it is sufficient for the feelings to be present only intermittently in order for us to have the emotion.[20] An explanation along the lines of the feelings being dispositional has the attraction of allowing us to explain how it is the feelings can erupt from time to time, whilst being apparently absent for long periods, an account which corresponds to the description of a typical state of enduring anger which I described above. We could argue that the initial stimulus arouses the feelings in the first instance, and that until time or an apology has dissolved the anger, we are still liable to further arousal even at the thought of what had happened earlier. This provides an account which allows us to explain the potentially extended duration of emotions, as well as the absence of conscious feelings at all stages.

The problem of such an account is the conceptual gap between dispositional states and constant states. If my feelings are dispositions, then this suggests that during the periods between which my anger is triggered, I am not angry. Wollheim denies this in defending the position that if I am disposed to become angry if stimulated, then I am angry. But this surely reveals a wider problem with any account of enduring emotions which seeks to describe them as dispositions, in that we lose sight of the crucial distinction between

---

[20] Richard Wollheim favours the idea of emotions generally as dispositional states. See Chapter One of *The Emotions*. Peter Goldie holds that feelings need only be present on occasion during the emotion. See *The Emotions*, 69.

being disposed to anger towards someone, which indicates a constant *potential* for anger, and being angry with someone for a long period of time, which indicates a constant *state* of anger. There is a crucial difference between statements such as 'I get angry every time I see Margaret Thatcher on TV', and 'I've been angry with Margaret Thatcher ever since she took power'. The continuity of the emotional state in the second statement is lost if we treat this as a dispositional state.

Goldie offers a different (and I think better) option in which he distinguishes between emotions and what he calls 'emotional episodes'. In the account of anger I offered above, the episodes would be the moments when the feelings come to the forefront of the mind, and we are, as it were, consumed by the emotion. Goldie states his position as follows:

> [I] think is it reasonable to say that, without at least episodes of such feeling, of which you can be more or less aware, an experience cannot be an emotional one. But once the distinction between an emotion and an emotional episode is clear, I do not need to insist, for example, that my enduring love for her (the emotion) need involve feelings at all times, even when I am playing football, although it will be true that my emotion at that time involves various dispositions to have loving thoughts and feelings towards her. So, looking back, I can truly say 'I have loved her for twenty years, without ceasing to do so for a second'.[21]

The attraction of such an account is the reconciliation of the duration of the emotion with the absence of feelings over the same period of time, whilst overcoming the problems encountered with Wollheim, where we struggle to distinguish between purely dispositional states and constant ones. When we describe someone as a jealous type, then this is different to describing him as someone who has been jealous of his brother for years. The first usage indicates that this is a person who is more likely than most to become jealous on seeing his wife chatting with another man, of a colleague who gets a promotion, or a neighbour with a faster car. This indicates that this person will tend to have many different and distinct experiences of jealousy. But this needn't exhaust what we can do when we attribute a disposition to someone with regard to their emotions. With the case of the man who is jealous of his brother, we can understand this to mean

---

[21] Peter Goldie, *The Emotions*, 69

that he tends to construe situations involving his brother as ones in which he wishes he could be. The first case involves a disposition to react in a consistent way to a wide range of different stimuli, whilst the second involves a disposition to react in a consistent way to the same stimulus. When we talk of the dispositions involved in loving someone for years it is surely dispositions of this second category which are involved. The consistency in the way we react to a particular individual does the work of justifying the claim that it is the same emotion.

This line of argument can be supplemented by taking a slightly more liberal approach to the role of feelings within emotions than the one which suggests feelings must always be present. In many cases, our ongoing emotions may well guide our behaviour even when we are not fully aware of feelings. When deciding who to exclude from the Christmas list, we might just unthinkingly cross off the brother with whom we argued last Christmas. Similarly, we might see the person we love come home exhausted, and just prepare a meal and open a bottle of wine. Neither of these acts need be accompanied by any surge of feeling, yet the actions suggest a way of construing situations which reveal feelings that condition our responses towards that person. When such behaviour is consistent over time, and there are periodic moments of greater awareness of our feelings, we describe such cases as a single emotion. Such a view allows us to make room for feelings waxing and waning, and of our being sometimes unaware of them, but retains the picture of a coherence and consistency in our emotional life which is true to our phenomenology. Taken together, it is the consistent behaviour which we take to reveal our feelings by revealing the way in which we construe the situation, combined with the periodic awareness of those feelings, which will justify the claim of having experienced the emotion.[22]

There will still be difficult borderlines cases, such as those where we hear statements like 'I loved her but I didn't realise it until she was gone, and I know I didn't behave like it at the time', but these just *are* complex cases, and any satisfactory account of our emotions must make room for the sort of distortion or clarification of our past feelings which may occur. But on balance, the justification of the claim to have experienced an emotion

---

[22] Perhaps the most famous discussion of this phenomenon is Hume's discussion of the 'calm passions' in section IV of Book II of the *Treatise*. His description of a passion becoming 'settled principle of action' is similar to the line I shall be developing in the next chapter with regard to how emotions can guide our behaviour without being constantly at the forefront of our minds.

over long periods will be one based on consistent behaviour, and feelings of sufficient strength and duration, that taken together they come to characterise the way we view the period in question.

There is one further argument which can be offered in favour of this view, to the effect that there are other areas of our phenomenology which don't require us to be in a constant state in order to attribute constancy to us. Gilbert Ryle makes an analogous point when attacking epistemologists who argue that believing or knowing are 'one-pattern intellectual processes in which these cognitive dispositions are actualised.'[23] As Ryle points out, even if people constantly repeated what they claim to know, this would not be enough to satisfy us that they really did know something. We would have to see them inferring, imagining, saying and doing things in order to be convinced. Similarly with emotions, we don't have to see people constantly affirming they are angry to decide that they are. Just as we can attribute to people the belief that the earth is round without them constantly affirming it, we can attribute an emotion to someone without their constantly feeling it. The attribution of an emotion to someone depends of a range of criteria, of which awareness of feelings is only one.

'Feeling Towards'

Many of the arguments set out so far in this chapter have referred to the concept of 'feeling towards' without setting out precisely what is meant by it. As this will play a crucial role in the work to come, it is worth taking some time to address this point now. In the first instance, it appears odd to describe a feeling as a perception of the inner state of one's body, indicating that the feeling is something internal to us, and then to talk in terms of 'feeling towards' someone else - someone external to us. In trying to explain this concept, I shall closely follow the ideas of Goldie[24] in setting out how our feelings can be said to have an object, and thereby explaining how it is we can talk of feeling something 'about the world', as against feeling something within our body.

Goldie's first move is to acknowledge that our feelings lack any direct intentionality. Nor can feelings, as I have already discussed, reveal what a specific emotion is. Instead, he

---

[23] Gilbert Ryle, *The Concept of Mind*, 44-45.
[24] Peter Goldie, *The Emotions*, 58-62.

borrows a term from William James, and argues that we must understand feelings as being 'combined in consciousness' with the object of an emotion. There is an association of ideas which connect our thumping heart with the charging elephant, providing our feeling with a 'borrowed intentionality'. This means that the object of our feeling is no longer a psychological or bodily state, but a feature of the external world about which we now have 'feelings towards', which is captured in the concept of 'thinking of with feeling'.

This conceptualisation can be brought out when we consider certain paradigmatic examples of coming to think of things differently because the thought is now infused with feeling in the way I described earlier with the example of desiring to see *Madame Butterfly*.[25] Take the example of the child who has been told that fire is dangerous, and has accepted this as true, but whose curiosity nevertheless leads her to plunge her hand into the flames. The upshot will be a way of thinking about the danger of fire which will differ from that which preceded the experience.[26] The sensation of the searing pain, the feeling of shock, and the association of these phenomena with the fire will enliven the thought of the danger in comparison to the more coldly formal knowledge the child held previously. The content of the thought will be different, and this difference is that the thought is now held with a *feeling* of fear, as against a feelingless belief about danger.

A further analogy is the difference between when a colour blind person[27] with a reliable guide utters the words 'That rose is red', and when a normally sighted person utters the same words. The language is the same, but the concept each person uses is different in the two cases, because the normally sighted person is seeing the rose as red.[28] A similar distinction holds between thinking of something as being dangerous in a cold, formal sense, and thinking of it as being dangerous in feelingful sense – one which implies fear.

If these arguments are correct, then we have the conceptual framework in place to combine the advances of the cognitivist theories which I discussed in the last chapter with the arguments to the effect that emotions must contain feelings. The conclusion which emerges from these thoughts is that we have found the means to account for the

---

[25] See page 38 above.
[26] Cf Stocker, *Valuing Emotions*, 47.
[27] I take this example from Peter Goldie, *The Emotions*, 60.

intentionality and the cognitive content of the emotions on the one hand, with the feelings on the other. The beliefs and desires which partly constitute our emotions are, moreover, not distinct elements from the feelings, but suffused with them in such a way that the beliefs and desires are filled with the feelings which make the overall experience distinctly that of having an emotion. It is in this way that emotions can contain beliefs and desires without ever being reducible to them.

There are, however, a number of other central claims which I have made here which will be part of the foundation of what I shall be arguing for when I come to look at the relationship between emotion and moral value. The most important of these is that our feelings towards the world represent a stable, and often subtle backdrop to our general perspective. I have tried to show that we often have only limited awareness of our feelings, and that the existence of stable feelings is the most convincing way of explaining the stability of a wide range of attitudes and commitments which are often retained over extended periods. I shall be arguing that this 'background of directed feeling' is what provides the foundation of judgements of value, and it is when these feelings erupt into the open in the form of emotion that we become fully aware of precisely which directions our feelings run in. The purpose of the next chapter will be to flesh out exactly how it is that our feelings acquire a particular direction in the first place.

---

[28] For a full account of this sort of 'concept dualism' and the use/mention distinction, see David Papineau, *Thinking About Consciousness*.

3. Belief and the Direction of Feeling

Introduction

I concluded the last chapter by saying that we have a background of directed feeling which explains the consistency and coherence we have in our emotional lives, as well as in our personal commitments and projects. I also claimed that it is when the emotions are aroused that the precise direction and character of this feeling is revealed. The aim of this chapter will be to elucidate this claim by setting out a conceptual scheme which describes how it is that feelings come to be directed in one way rather than another. The main thrust will be that we can understand our emotional development in ways broadly analogous to the development of our thought, and I shall be adapting the ideas of Donald Davidson with a view to establishing this. I shall then go on to support the conceptual framework laid out here by drawing on some relevant findings in child psychology. What will emerge is far from a Davidsonian theory of the emotions[1], but I shall nevertheless try to show that his theory of how thought develops can be usefully co-opted into the project of understanding how emotions develop.

I shall also be revisiting the issue of what is meant by saying that emotions are partly-constituted by beliefs or perceptions. If the material from child psychology is to be accepted, then we can attribute emotions to infants who are only a few weeks old. If this is the case then it follows that if we hold that emotions require beliefs, and that it is possible to hold beliefs prior to acquiring a language, this would mean that beliefs need not have propositional content. This areas of the thesis therefore requires elucidation, and I shall argue that the intentional states which partly constitute emotions must be located across a broad spectrum, starting from simple representations which contribute to an infant fearing an unfamiliar person, to more complicated construals such as those involved in a company executive fearing the possible outcome of a restructuring of the firm. It is in allowing this broad range of intentional states that we can attribute emotions to both infants and certain animals whom we might wish to describe as frightened or sad, as well as to sophisticated adult humans.

---

[1] I have already criticised Davidson's specific ideas in this area. See the chapter 'Emotion and Cognition' above.

A further key issue I shall be looking at is the unusual way in which a broad degree of similarity in emotional responses across different cultures appears to co-exist with a wide degree of diversity in the object of emotions. Research from within social anthropology and evolutionary theory appears to show that people from cultures radically different to our own often respond emotionally in the same way we might judge appropriate ourselves, but react very differently in other cases. The emotional lives of those in different cultures appears to vary more widely than our physical make-up, but less widely than language. In order to address this I shall look at the social constructionist viewpoint, and I shall try to show that we need a picture of our emotional development in which we start with certain capacities, naturally occurring in humans, which are developmentally open, such that different cultural norms will result in people's emotions being shaped in different ways. This will be of crucial importance in later chapters when it comes to explaining variations in moral outlooks within and across cultures, and the arguments set out here will complete the groundwork for the discussion of the relationship between emotions and moral value which is to follow. But in the first instance it will be helpful to try and clarify how it is that we come to feel towards different parts of the world in the way we do.

Triangulation

What is required is an understanding of how it is that we come to feel a particular way towards some objects, and differently towards others. In certain cases, the features of the object in question can provide much of the explanatory force, as in the example William James offered of an elephant charging towards us.[2] The properties of the charging elephant appear to make it ideally suited to arousing fear when combined with certain apparently straightforward perceptions. A relatively crude understanding of the possible impact of an elephant and its direction of travel would be sufficient to arouse feelings of fear. Even this assumption will turn out to be more complex than one might at first think, but we need a much wider theory than this in any case. If we accept that there is nothing inherently more dangerous about a spider than a cat, then why do many of us fear the former rather than the latter? We need an account of how it is that objects in the world come to be seen as bearing evaluative properties which can provoke

---

[2] William James, 'What is an Emotion?', *Mind* 9, 191.

particular emotions – to understand how it is that a spider can come to be seen as fearful and a cat as cute.

The model which I shall suggest to explain this is based upon Davidson's concept of triangulation, which he uses to explain the emergence of thought.[3] Davidson describes his basic thesis as follows;

> The basic situation is one that involves two or more creatures simultaneously in interaction with each other and the world they share; it is what I call *triangulation*. It is the result of a threefold interaction, an interaction which is twofold from the point of view of each of the two agents: each is interacting simultaneously with the world and with the other agent. To put this in a slightly different way, each creature learns to correlate the reactions of other creatures with changes or objects in the world to which it also reacts.[4]

Once the initial correlations have been established, then each of the subjects is in a position to expect the other to respond in a certain way towards the relevant stimulus. In the case of a table, an infant will expect a parent to use the world 'table' when referring to a particular class of objects, and will expect certain responses when she comes to use the word herself. The correlations are set up in the first place by repeated, consistent responses to the same stimulus, and when the child first experiments with the response herself she is rewarded, thereby cementing the link.[5] All of this requires certain fundamental similarities between us. "To understand the speech of another, I must be able to think of the same things she does; I must share her world."[6]

The question now becomes to what extent this serves as a useful model for understanding our emotional development. One initial objection might be that whereas we can point to a range of relevant physical objects and say the word 'red', there are no physical objects which provide us with instances of 'fearful' where the physical objects carry quite so much of the burden of explaining how we come to understand what the concept means. Surely part of the account of the perception of redness will come in the

---

[3] The material I shall be drawing on comes from essays six to nine in *Subjective, Intersubjective, Objective.*
[4] *Ibid.* 128.
[5] *Ibid.* 86, 117.

form of the physical properties of light-sensitive cones in the retina, and their tendency to be excited by the different light wavelengths? What analogous properties can we point to in spiders to explain their tendency to arouse fear, which we would point to in the Chinese flag to explain its tendency to elicit the experience of seeing red? The spider seems to lack the properties which elicit fear in the direct way in which a flag can elicit the seeing of red. There is a sort of shapelessness to the class of fearful objects, insofar as we cannot pinpoint the sort of common properties which are easier to identify in the class of red objects.

An obvious initial response is that this is scarcely an issue confined to explaining our emotional development. When we learn our language, we may well start by associating 'mama' and 'dada' with certain specific spatiotemporal objects, but we eventually break through into speaking of a world of unobservables, abstract objects, imaginary objects, and so on.[7] But to offer this response is already to concede more than one needs to in order to defend the relevance of triangulation. Daniel Hutto has pointed out that there are considerable problems in arguing for mind-independent features of the world which correlate neatly with the colours we perceive.[8] I do not wish to go into the technicalities of this debate, but simply to highlight the pitfalls of assuming ostensive definition of colours provides a clear-cut contrast to learning what is fearful. Whilst it may be true that the objects different people find fearful may lack the obvious core of essential properties that unite the objects seen as red, we need a much richer theory of how we learn concepts than one which runs along the lines that we simply learn what red is by having someone point to an example and say 'red'. How do we know exactly which property is being picked out? How do we know how to identify objects as red in cases where no-one is pointing it out to us? How do we learn that different shades can count as being shades of the same colour? This is one of the issues raised by Wittgenstein in his discussion of what happens when we teach people concepts by means of examples. After a while, we reach the stage where we give examples and there is an 'and so on' which 'points beyond' the examples. The way in which this works cannot be explained

---

[6] *Ibid*, 105.

[7] For a fuller discussion of this, see W V Quine, 'Speaking of Objects', in *Ontological Relativity and Other Essays*.

[8] Daniel Hutto, 'The World is not Enough: Shared Emotions and Other Minds' in Peter Goldie, ed, *Understanding Emotions: Mind and Morals*.

by the pointing and saying the word at the same time, as it is precisely this beyond to which we are now pointing.[9]

One way forward emerges when we consider one of the components in Davidson's discussion of triangulation – reward. A crucial factor in how it is that the child comes to correlate successfully words with objects is when the other person in the triangle rewards her with praise on hearing the appropriate term. But how is it that the child perceives this as reward rather than punishment or indifference? The reward is claimed through recognition of certain facial expressions and tone of voice, as well as behavioural responses characteristic of praise. The reward takes the form of excitement in response to these perceived messages.[10]

Hutto describes this as 'the intersubjective dimension of experiential concept learning'.[11] A crucial part of what is going on in the triangle concerns the behaviour on the part of the carer and the child, and the impact this has on the child. For example, both must focus on the target object and cross-check that the other is doing the same. The response involved will be both 'intentional and experiential', in that we must not only focus on a specific object, but experience it in roughly the same way – both teacher and learner must focus on a red object, and both must see it as red. Were this not the case, it would be impossible to triangulate – I would never know if you were seeing the same object as me. We must assume a 'broad common ground' between us in order to learn common concepts of experience and objects.

An example of this may help to bring out the key points. The child approaches the fire, and the carer's face is transformed in a particular way, just as the cry of 'Stop!' is uttered in a particular tone. The event is repeated on several occasions with the same response, and the unpleasant feelings elicited in the child by the carer are ones which the child comes to associate with the fire. She comes to 'feel towards' the fire in a particular way under direction from the carer. - she comes to fear it. As she grows older, she learns that the range of objects she has been taught to fear are called dangerous. The triangulation comes in the form of the child and the carer each being aware that the other is switching

---

[9] Ludwig Wittgenstein, *Philosophical Investigations*, § 208.
[10] Cf *ibid*, "I do it, he does it after me; and I influence him by expressions of agreement, rejection, expectation, encouragement. I let him go on his way, or hold him back; and so on."
[11] *ibid*

attention back and forth between the other person and the fire. The child starts with a capacity to perceive and respond to certain stimuli (in this case the expression of concern), as well as the ability to associate the concern with the fire. The result is that the child avoids the fire. The ability to learn in this way is part of what it is to share the same world.

This schema is necessarily broad, but it is only intended to capture the fundamental process through which we come to see objects or events as fearful, disgusting, worthy of anger, and so on. The process is ongoing, played out through numerous transactions, year after year, between child and carer. This continuous process results in the acquisition of a background of directed feeling which orients us towards the world. After a while, we no longer need the constant input of the carer, because we have now come to feel a particular way towards particular states, objects and activities. The learning process inculcates habit, which means that lessons learned can continue to guide us without the carer present.

It also provides us with a starting point for understanding why it is that we have a degree of consistency in the emotional lives of people across different cultures, coupled with a wide degree of diversity. I have suggested that all of us have the capacity to express and perceive feelings in the face and voice of another, and that there will be certain objects which are more suited as candidates for certain emotions rather than others. To take an obvious example, the consistently harmful potential of fire, coupled with the tendency of parents to care about the welfare of their children will result in the greater likelihood of children across different cultures learning to fear fire though the process I have suggested this also leaves space for a wide degree of variation in what we learn to feel both within and across different cultures, which reflects the more culturally specific nature of certain fears which we have. We can imagine entire generations of English children being brought up to fear the abominable snowman, whilst none of their Eskimo counterparts are even taught what this is. Similarly, one might have either a wholly Christian culture with different denominations, some of whom teach children a deep-seated fear of the cross, whilst others teach them to associate it with love and happiness. Socio-cultural norms thereby come to be transmitted through the education of the emotions.

Many of the ideas set out here can be given direct support from research carried out in the field of child psychology.[12] Specialists in this field describe a process called 'social referencing', in which infants respond directly to the facial expressions and tone of voice of their caregiver, by treating objects with greater wariness or enthusiasm depending on the cue they receive. When presented with an unfamiliar object, they will often turn to the caregiver, whose reaction will shape the response of the infant.

A further aspect of infant emotional behaviour is the habit of comforting or hurting others. Babies of up to twelve months of age may become distressed on perceiving the distress of someone else, and over the following year many will develop the habit of actively trying to comfort or even protect a victim. A converse pattern of behaviour is evident when it comes to hurting others. If a baby has caused distress in a younger sibling, the most common response is to hurt them even more, and many infants even use strategies cunningly designed to distress the sibling. Perhaps unsurprisingly, the babies most likely to cause deliberate distress are least likely to comfort, but why should some people come to be moved by the distress of others whilst some deliberately cause it?

In cases where one child has hurt another and the mother explains that this is wrong and comforts the victim herself, the wrongdoer will tend to develop into a comforter. In cases where children have not been exposed to this sort of example, and especially where they have been victims of abuse, they will tend to become either indifferent to the perception of distress in others, or even respond with hostility. It seems that what we might conventionally term a sympathetic response is one which we learn, and various factors will influence the extent to which we develop this characteristic. For example, children whose parents talk about their feelings will go on to be more adept at describing not only their own emotions, but recognising emotions in others.[13]

In short we have certain shared abilities which are open to being developed in different directions, and are therefore susceptible to cultural variation. The direction which our feelings are given will come in the shape of massive input from those around us, which

---

[12] For a full summary and discussion of the research in this area, see Paul L Harris, *Children and Emotion*.

[13] See Dunn, Brown and Beardsall, 'Family Talk about Feeling States and Children's Later Understanding of Others' Emotions', in Jenkins, Oatley and Stein eds, *Human Emotions: A Reader*.

will follow set patterns. We will tend to be discouraged from approaching fire, and (one would hope) encouraged to be sympathetic to siblings, praised for what is taken to be good behaviour. In this way we come to have what I have described as a background of directed feeling. The input we receive covers much of the world around us, which is now enlivened by the feelings we have towards it. Certain classes of objects are perceived with fear, some with disgust, others with delight. The world is now 'alive' as a result of the feeling-laden beliefs and perceptions which we have acquired.

This raises the question as to whether or not our emotions might not be understood entirely as cultural phenomena, rather than the social development of certain shared characteristics, as I have suggested. It is this aspect of our emotional development which has provided the focus for one of the approaches to the subject over the last twenty years.

Social Constructionism

The social constructionist[14] theory of the emotions claims that 'emotions are characterised by attitudes such as beliefs, judgements and desires, the contents of which are not natural, but are determined by the systems of cultural belief, value and moral value of particular communities'.[15] The social constructionist begins from two points, the first of which is the intentionality of the emotions, which I have already discussed.[16] The second point is the diversity of emotions which one finds across different cultures, which is present in two different forms. Firstly, people in different cultures have different emotional responses to the same stimulus. Japanese people tend to feel disgust at someone sneezing in public, and shame at the thought of doing it themselves, whereas British people would tend to be rather more sympathetic or indifferent. Secondly, there is research which suggests that different cultures develop distinct emotions. To stay for a moment with the Japanese, they are said to experience an emotion of *amae*,[17] which is argued to be untranslatable, but which broadly corresponds to presuming upon another's kindness or basking in another's indulgence. *Amae* is attributed both positively and negatively, and examples of when it might be felt would be when a young man is looked

---

[14] See *The Social Construction of the Emotions*, ed. Rom Harr—.
[15] Claire Armon-Jones, 'The Thesis of Constructionism', *ibid* 32-57.
[16] See 'Emotion and Cognition' above.

after by his mother prior to marrying, or even in teacher-pupil relations. The range of different circumstances when the Japanese are said to experience *amae*, and the moments at which it is considered appropriate or inappropriate appear to indicate that it is quite distinct from any emotion which people from other cultures have.

The social constructionist argues that the intentionality of the emotions coupled with their cultural diversity supports the claim that emotions are culturally constructed. The theory allows for some universal emotions of a particularly crude physiological character, but the cultural production of emotion means that 'the bulk of mankind live within systems of thought and feeling that bear little but superficial resemblances to one another.'[18] The acquisition of emotions entails learning cultural norms, standards and principles, and an understanding of what is appropriate in which circumstances. Emotions themselves are to be explained in terms of their functional role in the maintenance of the community, a role they accomplish by inducing individuals to refrain from what is deemed undesirable behaviour and promoting desirable actions and attitudes.[19]

Before considering some responses to this position, it is perhaps worth trying to specify more precisely what a socially constructed emotion would be. Harr— warns against what he calls the 'ontological illusion' of assuming that there is something there, of which the emotion word is a mere representation. He argues that it is wrong to be misled by physiological changes in the body, which are simply 'incidental effects' of the emotion. The emotion itself has to be understood as existing 'only in the reciprocal exchanges of a social encounter'.[20] But perhaps the key comment which enables to see exactly how the social contructionist understands what an emotion is comes when Harr— states that the way to avoid the ontological illusion is opened up by the 'linguistic turn'.[21] It seems that emotions are to be understood as deriving their existence from within our language and our exchanges with other members of our community with whom we develop our understanding of how the world is. Emotions can be seen in the same light as values and

---

[17] See H Morsbach and W J Tyler, 'A Japanese Emotion: Amae', in *The Social Construction of the Emotions*, ed. Rom Harr—, 289-307.
[18] See Rom Harr—, 'An Outline of the Social Constructionist Viewpoint', *ibid*, 2-15.
[19] Claire Armon-Jones, 'The Thesis of Constructionism', *ibid*, 33-34.
[20] *Ibid*, 5

beliefs which bind communities together, and which derive their meaning from the socio-linguistic practices out of which they have emerged. Once we recognise this, we are immune from the illusion of there being a place beyond language where we find the real emotion.

Rather than dealing with the issues surrounding the linguistic turn generally, I shall focus on social constructionism's specific contribution to theory of the emotions. The central problem with the thesis seems to be, as Peter Goldie has phrased it, that it represents a 'significant overstatement'.[22] To begin with, one can accept the premises of the argument which centre on massive cultural variation, without this justifying the conclusion of there being no natural basis to our emotions, in that there can be natural traits which are shaped by environmental circumstances. Our learned eating habits in early years can have a considerable impact on the way in which we develop physically, yet would not want to say that the body is entirely socially constructed. Two further counter-arguments also seem to pose problems for the social constructionist position. The first is that there is evidence to suggest that the degree of cultural diversity in the emotions is more limited than one would expect if emotions were developed exclusively within individual cultures. There is, for example, research which appears to prove that the facial expressions related to the expression of particular emotions is similar even between cultures where we are aware of no way in which they could have influenced one another.[23] If the emotions were as influenced by culture as the constructionist would have us believe, there is no reason why the expressions of emotion in different cultures should resemble each other any more than languages do.

A second argument which weighs heavily against the constructionist thesis is the role of bodily feeling in emotion. The problem this poses is expressed by Claire Armon-Jones:

> According to Constructionism, a socioculturally constituted emotion is an acquired response. This requires that the elements constitutive of the emotion are ones which are capable of being acquired by the agent. Consequently, it is

---

[21] *Ibid.* For the clarion call of the linguistic turn, see Richard Rorty, 'Metaphilosophical Difficulties of Linguistic Philosophy', and for a brief summary of his reasons for moderating his position, see 'Tweny-five Years Later', both in Richard Rorty ed. *The Linguistic Turn.*

[22] Peter Goldie, *The Emotions*, 92.

[23] See P Ekman and W V Friesen, 'Constants across Culture in the Face and Emotion', reprinted in Jenkins, Oatley and Stein ed. *Human Emotions: A Reader*, 63-73.

essential to constructionism that an account of emotion be given in which emotions are neither identifiable with, nor have the same ontological status as, phenomena such as sensation and perception. This is so because, although perceptual skills, for instance can be acquired, we also have reason to believe that perception and sensation are not essentially skills acquired by training but are natural phenomena which exist prior to the acquisition of any sociocultural frame of reference within which we might want to explain them.[24]

The problem then becomes how one deals with the feeling of my body tensing up as I see a car mounting the pavement and heading towards me, and an African tribesman experiencing the same feeling as he perceives a tiger racing towards him. If one accepts, as Armon-Jones does, that feeling is part of emotion, and that physical tension is an example of a bodily feeling, then there appears to be reason to believe that the emotions of people from radically different cultures have some sort of shared natural dimension to them. Armon-Jones' response is to resort to what she sees as a Wittgensteinian position in which she argues that feelings can be understood as 'relevant emotion attitudes'.[25] But this seems like a desperate move to avoid the choice which the constructionist must face: she must either exclude feelings from the concept of emotion and thereby confront all the problems faced by cognitivist accounts, or else include feelings and allow in the common physiological features of emotions which suggests the kind of 'naturalism' she has spurned in favour of social construction.

Where I think the constructionist thesis goes wrong is with its approach to intentionality. I have argued that emotions must have an object, but the constructionist wishes to go one stage further and argue that an emotion is defined by the *specific* objects it has. This means that fear is not partly constituted simply by having an object, but by having a class of specific objects which arouse it. As such, the class of objects which members of one community fear would define what fear is there, and the class of objects which people fear in a different community would define what is to count as fear there. In that it is logically possible that the two classes may share none of the same objects, one can then conclude that use of the concept 'fear' to denote a single emotion across all cultures is

---

[24] Claire Armon-Jones, 'The Thesis of Constructionism' 43, in Rom Harr— ed, *The Social Construction of Emotions*.
[25] *Idid*, 48

misleading.[26] It is this view which opens the door to an emphasis on the sociocultural influences, and which leads to the conclusion that fear of the stock market falling is seen as a different emotion to fear of censure by the shaman, as against the same emotion with a different object. But even if these arguments expose deep flaws in the constructionist thesis, they also bring to the fore precisely the issues which this thesis seeks to highlight. The cross-cultural similarities in the expression of emotions, and the common tendency towards bodily feelings cannot obscure the fact that the cultural variations which the constructionist has focused on require an explanation which involves an apparently natural capacity which is open to significant cultural influence. It seems we need an explanatory framework which falls between those we might use to explain a universal physiological feature such as a reflex on the one hand, and a cultural phenomenon such as a particular language on the other. I suggest that triangulation provides us with just such a framework.

Belief and Emotion

One of the conceptual issues which emerges from the overall account I have so far provided is how we are to understand what a belief is in the context of our emotional lives. I have argued throughout that our emotions are complex phenomena partly constituted by beliefs or perceptions, and this is brought out when we try to explain something like our fear of our football team being relegated. Such a fear will be characterised by beliefs about the league tables and the quality of our players which are held with feeling. But this raises clear problems with regard to attributing emotions to infants or, as we might well do, to animals. If I snatch the dummy out of my baby's mouth, and he then contorts his face in a certain way, starts crying, and violently waves his arms and legs up and down, we would describe this as anger. If I leave the house, and I hear my faithful old dog scratching at the door, and then whimpering as I walk down the garden path, it is equally appropriate to describe him as sad.[27]

I am inclined to say that the attribution of emotion in both cases is appropriate, but this requires clarification of what concept of belief is being used in the overall account of

---

[26] Perhaps the quickest and easiest response to this line of argument is simply to point out that there is a single class of the objects of fear, and that class contains all fearful objects.

[27] For detailed discussions of emotions in both animals and infants, see Martha Nussbaum's *Upheavals of Thought*, chapters two and four.

what emotions are. If emotions are partly-constituted by beliefs, and certain non-human animals and pre-linguistic infants can experience emotions, then it cannot be the case that beliefs must have propositional content. Even the most ardent dog-lover would surely not say that the dog thinks to himself, 'my master is leaving'. This is a separate challenge to the one I examined earlier,[28] in that the question no longer turns on whether or not emotions can be reduced to propositions, but on whether or not the beliefs which partly-constitute emotions require propositional content. One could hold this, as well as believing that feelings must also be present.

This debate is vast and complex, and there is not space here to do it full justice, but I shall set out some of the objections to the concept of a belief which my thesis commits me to, and then offer a characterisation of what belief is which both meets the needs of this thesis and attempts to explicate our ascription of belief in relevant cases. The difficulty for those who would argue against the possibility of thought without language comes in the celebrated example offered by Norman Malcolm.[29] Imagine your dog is chasing the neighbour's cat, which races towards a convenient oak tree, but, unseen by the dog, swerves at the last moment and races up a nearby maple. The dog runs up to the oak, rears up, paws at it, and barks excitedly at the branches above. Surely we could say the dog is barking up the wrong tree because he thinks the cat is up there.

Davidson's response to this is to try to throw into doubt our intuitive practice of attributing beliefs and desires in such cases.[30] If we say the dog believed the cat was up the oak tree, could we say that the dog believed the cat went up the oldest oak tree in the garden? If not, then how can we distinguish between quite different things which the dog might have believed, some of which might be true, and others false? It seems we need some sort of *de re* description which suits the dog in this scenario, such as the dog thought 'the cat ran up that particular tree'. But this simply isn't possible unless we assume implausibly that the dog has other beliefs about trees. This objection obviously springs from Davidson's wider theory of what it required for us to have a thought – 'We identify thoughts, distinguish among them, describe them for what they are, only as they

---

[28] See 'Emotion and Cognition' above.
[29] Norman Malcolm, 'Thoughtless Brutes', in *Proceedings and Address of the American Philosophical Association, 46 (1972-73).*
[30] See Donald Davidson, 'Rational Animals', in *Subjective, Intersubjective, Objective*, 95-107

can be located within a dense network of related belief"[31] In addition, each thought depends on a sentence through which it is expressed. Such a view means that if we start ascribing individual beliefs to dogs or infants, then we find ourselves having to explain how they could have the many others required for the first to make sense, as well as attributing language to them.

Daniel Dennett has offered one immediate response to what he sees as the practice of requiring all creatures to have available to them the sort of fine-grained distinctions in thought which we can make through language and the propositional content it provides us with.[32] Dennett argues that we don't need to know exactly the way in which an agent conceives of his task – we only need a rough knowledge of what the object of concern is, such as food or cats. Human language just cuts too fine for its distinctions to be available to a dog, but the dog nevertheless has its own ways of discriminating things which we can learn about if we assume the 'intentional stance' (to which I shall return to a moment). In short, one can argue that we just don't need the degree of sophistication provided by a complex language. If non-human creatures can be seen to be discriminating in a systematic way between different objects, and behaving systematically in response then this could be taken as evidence for beliefs about those objects.

A further issue we encounter if we follow Davidson's line is that we are required to explain the transition which we apparently make from having no thought or language, to having a 'dense network' of beliefs, desires and intentions. Davidson recognises the difficulty here, and questions whether or not we might ever be able to give a satisfactory description of 'half-formed minds'. He nevertheless attempts to offer some ideas on the 'prelinguistic, precognitive' states which must precede and prepare the ground for thought, and it is here that he introduces the concept of triangulation.[33] This means that if there is to be a convincing alternative to Davidson's view of belief as necessarily being an attitude towards a proposition, it is perhaps to be found in a state which we might agree with Davidson to be prelinguistic, but not necessarily precognitive.

---

[31] *Ibid.* 98.
[32] See Daniel Dennett, *Kinds of Minds*, Chapter Two.
[33] See 56-62 above.

This is, broadly speaking, the position taken up by Jonathan Bennett.[34] Bennett offers only a partial characterisation of what a non-linguistic thought is, but his approach is more that if we use 'normal, reasonable, everyday standards', we can justify attributing beliefs to infants and certain animals. If we take the example of a bear loping across the ice towards what looks like a seal, but is really a rock, we can reasonably attribute thought in the following way. The bear has certain goals, such as eating food to stay alive. The bear 'registers' certain aspects of its environment, and systematically responds by running after, running away, watching, and so on. Under certain circumstances, the bear behaves in ways which indicate it has the capacity to makes errors, such as mistaking a rock for a seal. This means we can present a theory concerning the bear's goals, which is supported with reference to certain environmental stimuli such as seals and rocks which look like seals, and also by certain behaviour, such as running towards such objects and eating them when they are seals.

This means that the concept of 'registering' is anchored 'on the one side in epistemic input and on the other in behavioural output'.[35] Bennett emphasises that the conditions which are sufficient for us to attribute thought for 'everyday' purposes will not be *logically* sufficient for attributing thought, but he claims to have provided a coherent means of attributing thought to languageless creatures which at least meets the standards required of our common practice in this domain.

This position partially resembles the more prominent ideas set out by Dennett when he discusses 'the intentional stance'.[36] This is defined as '[t]he strategy of interpreting the behaviour of an entity (person, animal, artifact, whatever) by treating it *as if* it were a rational agent governed by its "choice" of action by a "consideration" of its "beliefs" and "desires"'.[37] The scare quotes are intended to alert us that we are using these terms in what might initially seem to be an inappropriate manner, in that these terms are not generally used with regard to single-cell organisms, vacuum cleaners, or ice-floes, all of which would fall under the usage set out here. Very briefly, this usage is defended with regard to animate objects on the grounds that our evolutionary heritage has provided us all, amoebas and humans alike, with survival and reproduction-driven goals with regard

---

[34] Jonathan Bennett, *Linguistic Practice*.
[35] *Ibid*, 33.
[36] I shall be looking at the presentation of this view as set out in Chapter Two of *Kinds of Minds*.
[37] *Ibid*, 35.

to which it makes sense to describe us as choosing, believing, desiring etc. to achieve the goals we aim at. Machines are given goals by humans, and therefore can also be meaningfully described as choosing certain options, as, for example, when a computer chooses to scan the hard disk when it has not been shut down properly. Within the context of this theory, thinking basically comes out as processing information for the purposes of achieving goals.

One factor which makes the debate particularly difficult is that there is a tendency for the different sides in the discussion to talk past each other. Davidson offers an account of what a belief is. Bennett offers a defence of attributing beliefs for the pragmatic purposes of predicting and understanding behaviour, and bases it on certain common practices. Dennett appears to use the same common sense approach as Bennett when ascribing beliefs, but then offers an account of what believing is which flies in the face of our common sense intuitions about human thought being distinctively rich and complex in comparison to the thought of other animals and machines.

When one comes to consider these rival arguments with regard to a theory of the emotions, we are faced with either letting too much in under the guise of belief, or else leaving too much out when it comes to emotions. If belief is simply 'registering' or 'processing information', then what separates human beings from guided missiles, where such terms can surely be used to describe what they do? If we require all beliefs to have propositional content, then how can we attribute emotions to infants and animals if those emotions contain beliefs?

In order to chart a passage between the conceptual rocks, we can start by observing that there is slightly less at stake when addressing this problem in the context of a theory of the emotions in comparison to what the discussion of what a belief is *tout court*. To focus for a moment on Bennett's view, we saw that one of the drawbacks comes in the form of our being able accurately to describe guided missiles as 'registering' objects and making predictions about their behaviour which includes reference to goals and the possibility of error. This conflicts with what I take to be the 'normal, reasonable, everyday' understanding of what believing is, which tells us that machines don't do it. But there is a safeguard within the theory of the emotions which prevents the problem from spreading in this way. We can surely avoid any move from machines being able to think,

to machines being able to have emotions, and this conclusion is blocked by stipulating that feelings are necessary constituents of any emotion. This means that even if we grant machines the power of registering features of the external world, this is insufficient to grant them an emotional life. Guided missiles may process information and make mistakes, but they certainly don't experience feelings of depression if they miss their target.

If this move is acceptable, then it reduces some of the pressure on how broad a concept we can permit to fulfil the role of the intentional state which partly constitutes any emotion. We can allow beliefs to be simpler in character than a Davidsonian would wish to allow, and yet avoid having to accept the more unwelcome consequences which may follow if we are concerned solely with beliefs. I accept that this simply side-steps the conceptual differences between Davidson and Bennett, but my aim is solely to provide a better understanding of the role of belief in emotion, not a full-blown account of what belief is.

The next step is to clarify what is required of belief in the overall theory of emotion which is being developed. This is best understood as the key to the intentionality of the emotions, in that it is our representations of objects and events which come to be feeling-laden. It is our beliefs which are 'combined in consciousness' with our feelings in such a way that we come to have 'feelings towards' certain features of the world. The question we can now ask is how complex these states must be in order for them to combine with feelings in such a way that they create what we would wish to call an emotion, and this will provide the outline for our understanding of the relevant intentional states in this context.

If we use our everyday ascription of emotion as a guide in answering this question, then we need a wide range of intentional states. In order to highlight this, consider the following two examples. A four-month old child is passed by his mother to another person. The child smiles broadly, waves its arms about, and issues a series of sounds which we generally take to indicate satisfaction. We can contrast this with a bond dealer staring at a screen, who smiles broadly and punches the air on reading a particular set of numbers and letters which flash before him. If we take both cases to be expressions of delight, then this seems to involve a simpler perception on the part of the child, and a

much more highly complex interpretation on the part of the bond dealer concerning what the figures and letters represent, and the likely consequences of this information. I suggest that the range of intentional states which we need to justify the attribution of emotions such as fear, disgust, happiness and so on to both infants and adults must be extremely broad if we are to justify this practice. In the case of infants and certain animals, relatively simple perceptions of changes in the immediate environment or their body must be conceptually sufficient to count as one such state. Conversely, the range and sophistication of our emotional life can increase immensely as we acquire language and the 'dense network' of beliefs, desires and intentions on which so much emphasis is placed by Davidson.

I suggest that the best way forward is to acknowledge that the intentional states involved in emotions must cover a broad spectrum. This includes simple perceptions on one end, which are recognitional states of objects in the immediate environment, to complex construals at the other end which involve sophisticated interpretation of symbols within a richly complex web of understanding. They will share a direction of fit, in that all beliefs will be representations of the way the world is, but they can be distinguished in terms of the way in which they represent their object. The simple recognitional states will have the 'target' of the perception as the object of the emotion, which is to say that the object of a baby's fear will tend to be the object she is looking at. With more complex beliefs, the target of perception may well be different from the object of the fear, for instance when the red numbers which a bond dealer perceives on the screen (the target of the perception) represent a potentially catastrophic fall in the value of his purchases (the object of his fear). Crucially, both the simple perceptions and the complex construals have the potential to be combined in consciousness with feelings, such that they can partly constitute an emotional state.

The best way to capture this broad spectrum of states which are involved in emotions is to say that feelings must be combined with either perceptions, which might be very simple recognitional states, or beliefs, which will takes us as far as complex imagined scenarios available only to those with a sophisticated language. This solution leaves open the question of exactly what a belief is and whether it must have propositional content, but this is a virtue rather than a vice in this context. It means that a Davidsonian could accept this line, and argue that animals and infants have emotions which involve

perceptions but not beliefs, but it also means that a Dennettian could equally accept it with the caveat that the creatures without language who experience emotions do so partly in virtue of their beliefs about the world.

When we say that an infant was sad when her mother left the room, we attribute to the infant the perception of the mother leaving the room which is combined in her consciousness with the perception of the state of her body. We don't assume language, but simply the capacity to discriminate between different states of affairs, and the capacity to have feelings which emerges as a result of perceiving things to be one way rather than another. When we come to describe the development of our emotional lives the acquisition of language will come to have a central role in helping us to explain cases such as the bond dealer. Language helps us to conceive of objects which are beyond simple sense perception, and to describe those objects in such ways that they can elicit feelings in us. It also helps us acquire a better understanding of the emotional lives of others by allowing us access to emotional states which can be more fully and precisely expressed through language than they could be merely through facial expressions or other forms of behaviour. This sort of approach is the one which I take Nussbaum to be arguing for when she claims that what we need is 'a multifaceted notion of cognitive interpretation or seeing-as, accompanied by a flexible notion of intentionality that allows us to ascribe to a creature more or less precise, vaguer or more demarcated, ways of intending an object and marking it as salient'.[38]

There is a further aspect of our emotional development which relates back to the earlier discussion of triangulation.[39] If the way in which we come to feel towards the world is indeed captured in this model, then the feeling-laden content of our perception of fire which we acquire through social referencing will colour the word 'fire' or 'cross' when we come to learn it. The experiments into social referencing start with infants of as young as twelve months of age, who respond 'with a fearful face' when perceiving a similar expression on the face of the mother.[40] If we accept the evidence of social referencing, then surely the more plausible model on which to understand how it is that emotions develop once children have the use of language is that previously inarticulate perceptions and emotional states are now expressible through language. This suggests that the world

---

[38] Martha Nussbaum, *Upheavals of Thought*, 129.
[39] See 56-62 above.
[40] See Paul L. Harris, *Children and Emotion*, 20-23.

is already 'alive' for us by the time we start to acquire language, which extends the range of our emotional sweep into the realms of the abstract and the physically remote, as well as deepening our understanding of others' emotions.[41]

The scheme I have set out here requires that our emotions initially involve simpler perceptions of the world, our understanding of which come to be enriched and extended once we acquire language. These simpler perceptions demonstrate a capacity to discriminate between different states of affairs, and form the basis of consistent responses to the same stimulus. I suggest this characterisation of the representational states present in emotions, and their relationship to language, provides an explanatory framework within which we can address two of the issues confronted in this chapter. Firstly, it helps to explain how we might justify our practice of attributing emotions more widely than to humans who have acquired language. Secondly, it elaborates on the view of emotions as natural capacities which are open to significant influence from our culture. On the picture painted here, we can explain why it is that infants will experience emotions from an early age, and then go on to deepen their emotional lives in adulthood as language enables them to express and understand better their own emotions and those of others.[42] It also helps to clarify how it is that the world comes to 'seem' one way to us rather than another, as we learn to respond in ways which reflect the standards of the wider culture.

I also think this would be compatible with a Davidsonian position, as even he acknowledges that when it comes to explaining what lies 'in between' prelinguistic mental states and thought, attributing intentions and desires is a means in which '[w]e have no better way to explain what they do.'[43] In this, I agree with him, and when taken with Bennett's arguments, it seems to present a reasonable case for justifying our attribution of emotions to some languageless creatures.

---

[41] Perhaps one way of bringing this out is to adapt one of the passages from Wittgenstein. 'How does a human being learn the meaning of the names of emotions (sensations in the original)?.....adults talk to him and teach him exclamations and, later, sentences. They teach the child new emotion behaviour.' Ludwig Wittgenstein, *Philosophical Investigations*, § 244.

[42] See Dunn, Brown and Beardsall, 'Family Talk about Feeling States and Children's Later Understanding of Others' Emotions', in Jenkins, Oatley and Stein edd, *Human Emotions: A Reader*

The conceptual scheme I have set out here represents the groundwork for the ethical theory which is to follow. I have suggested that our capacity for feeling towards the world is open to development in a range of different ways which come to given specific direction within our immediate environment and also by our wider culture. It is clear that this has significant bearing on our ethical life, given that there are distinctly ethical emotions such as shame and guilt. But there are major conceptual hurdles which still have to be crossed in order to move from the general comments on the existence of a background of directed feeling, to the more specific domain of the ethical. How do we come to distinguish between ethical and non-ethical feelings and emotions? How do we move from certain apparently primitively ethical emotions such as an infant being moved by the distress of another person, to moral concepts? It is to this which I shall now turn.

---

[43] Donald Davidson, 'The Emergence of Thought', in *Subjective, Intersubjective, Objective*, 128.

4. The Moral Perspective

Introduction

The aim of the first three chapters was to set out an account of the nature of emotion which will inform the moral theory which is to follow. The aim from now on is to defend the broad claim that our moral lives are a function of our emotional development, and more specifically that compassion and reason are jointly necessary and sufficient for what I shall describe as a proper morality. In the first instance, there is a need to demonstrate the relationship between our emotional development and the ethical domain. We need to come to a clearer view as to how it is that our emotions make morality possible, and what we mean by morality when we commit ourselves to this sort of sentimentalist position.

In light of this, the specific aim of this chapter will be to provide an account of what it is to have a moral perspective which follows from the earlier account of the nature of emotion. The key concept I shall employ in to explain the idea of a moral perspective is that of the 'internalised other', which I shall discuss in detail in the next section. This should be understood as the shaping of our personal identity by the needs of others, a process which I shall demonstrate is a part of our emotional development.

The Internalised Other

In order to establish what the moral perspective is, and how it develops, I shall draw on the material used by a number of writers with regard to the emergence of shame. My aim is not to provide an in-depth analysis of that emotion, but merely to use discussion of it as a vehicle to understand our moral development. By shame, I shall understand an emotion which is characterised by a number of different features. It is generally taken to be a response to states of character, such as shame at being a coward. It is usually indicates a sense of failure, particularly falling below a public standard. It often involves the sense of being revealed, and the typical response on the part of the person experiencing shame is concealment. The metaphorical image which is often used to express it is the gaze of another who looks on with disdain or disgust.[1]

We are forced towards the conceptual tool of the internalised other by a combination of elements which partly constitute our concept of shame. In the first place, an emotion which is a response to

---

[1] For detailed philosophical discussions of shame, see Bernard Williams, *Shame and Necessity*, Gabrielle Taylor, *Pride, Shame and Guilt*. For detailed discussion of shame from the perspective of psychology and social anthropology, see Paul Gilbert and Bernice Andrews eds, *Shame; Interpersonal Behaviour, Psychopathology and Culture*.

being revealed, indicates that there must be someone or something to whom we are revealed, a point which is a necessary condition for shame to develop in the first instance according to the empirical evidence to be canvassed below. If we tend to perceive, even if only metaphorically, the disdainful gaze, then there must someone who is metaphorically gazing upon us. And crucially, given that we can experience shame on our own, without the belief that the object of our shame is likely to be discovered, then a real external person cannot answer the conceptual requirements raised by these first two points. As such, we need the idea of a gaze which has been internalised, and has thereby become a part of who we are and a part of our understanding of both ourselves and the world around us, at least insofar as we are talking of instances of shame. The concept of the internalised other answers these conceptual needs.

It will useful to give a clearer account of what is to be meant by internalised other. By internalised, I shall mean that certain standards and ideals have become part of our identity. This means that when we come to evaluate an action or a state, we do so from a perspective which is partly constituted by standards and ideals which we have acquired from those around us, and which have become part of our way of understanding the world. As the term suggests, this means that these ideals are initially external to us, and come to be part of who we are. The 'otherness' indicates a contrast with what I shall label personal desires, which I shall use in a very broad sense to mean the purely self-serving goals we strive for. The contrast comes when the external ideals which have been incorporated into our perspective come into conflict with those desires. The 'other' is a metaphorical term for an influence which is initially that of a real person, but which comes to be part of us even when that person is no longer present. I do not wish to suggest that what we internalise becomes a rigid and unchanging aspect of who we are, and I shall indicate below how the content of the internalised other can alter. Rather than this, I am suggesting a source of influence which is initially external but which comes to shape the way we respond to the world around us.

Some care needs to be taken here. In particular, I am not suggesting a simple opposition between starting out with the internal and the selfish on the one hand, and acquiring what is initially external and morally ideal on the other. There are such things as external non-moral ideals. In cases where we see materialism as morally wrong, and we come enjoy the feel of purchasing endless items of designer clothing to the point when we simply can't go out unless we are covered in trendy labels, there is an obvious case to the effect that we have internalised certain non-moral ideals which we ourselves would acknowledge to be selfish ones. Nevertheless, the internalised other indicates the presence of at least some specifically moral ideals which contrast with personal desires. When we feel shame, we have often fallen short of the moral ideals which have become part of our perspective.

At this stage, let me pause to consider briefly two alternative explanatory accounts to the one I have offered. In the first place, what is the difference between the account I have offered and Freud's description of the super-ego? In Freud's account, similar in many respects to the one I have offered here, the aggressive tendencies of the child are directed back towards the ego. There it is taken over by a portion of the ego, 'which sets itself over and against the ego as super-ego, and which now, in the form of 'conscience', is ready to put into action against the ego the same harsh aggressiveness that the ego would have like to satisfy upon other, extraneous individuals.'[2] How does the internalised other differ from the super-ego? The most prominent way lies in the fact that Freud cannot be taken to offer a purely developmental account of how guilt or shame come into existence. His theory offers an entirely reductive view of such emotions, whose purpose is to inhibit aggressiveness so that civilisation can persist more effectively. Nothing in what I have said commits me to such a reductive theory of ethics, and it is difficult to see a theory as Freudian if it is denied ultimate recourse to certain non-moral basic drives such as eros or the death instinct. In practice, this means that Freudian shame and guilt are means to certain non-moral ends, whereas the internalised other has been introduced purely as an account of how wider moral standards become part of our identity. Moreover, the forms of behaviour which elicit responses in others are much wider than aggression. A wide range of different attitudes and behaviours will elicit responses which will shape our later attitudes and beliefs. A further distinction to the Freudian approach, one which will developed at length in chapter seven, is that I shall suggest that the content of the internalised other is subject to change through reason. We can learn facts that change what we feel, and we can reason about the nature of what we find shameful in ways which do much more than identify the causes of shame. In short, reason can do more than simply diagnose and reveal, it can also alter and improve, where this might mean altering the content of our beliefs to make them more internally coherent, or so that our perspective coheres more closely with the facts concerning the world around us.

A second possible objection is whether or not the internalised other could not be explained away as an imagined other. Rather than complicating the picture by arguing that our identity is altered, could we not offer a simpler account along the lines that constant exposure to someone who provides a model for our own behaviour, also creates an ongoing witness whose presence we imagine even when she is not there? This would mean that role of the internalised other could be adequately explained by the sort of thought present to us in the question, 'What would people say if they could see me now?'.

It is difficult to defeat this claim straight off, in that the possible range of circumstances in which the internalised other is present is co-extensive with the that of an imagined other. If it is a real

---

[2] Sigmund Freud, 'Origin of the Sense of Guilt', reprinted in *Guilt and Shame*, Herbert Morris ed, 54-58.

person who provides us with standards we come to internalise, then the first person whose standards influence us could always be the one we imagine. Even in circumstances where there is no actual possibility of their discovering my shameful state, it is logically possible for me to imagine their reaction if they did. The first response to this is simply to raise the question as to whether or not we really do imagine someone else looking on whenever we feel shame, and this just doesn't seem to be the case. We often feel shame without imagining anyone else seeing us, and even if the thought of what an important other would say may come to us, this surely represents an additional burden which the shame brings, rather than a necessary and sufficient condition.

A further argument against such an account comes when we turn to the empirical evidence of how our emotions come to be shaped by those around us. I have already indicated how the primary carer comes to influence the emotional development of a child,[3] and a similar process can explain how the ideals related to shame emerge. Such an account just fits better with the view that our identity is indeed altered, rather than with one which suggests we are simply susceptible to thoughts of what others think of us. An imagined other suggests the image of someone whose opinion might be of importance to us, but who is external. This just doesn't cut deep enough to describe the process of what occurs when our emotions are shaped. It is our identity which is altered and shaped through this process, and it is crucial to see the results as something which from a part of what we become - something internal to the self.

Perhaps the best way in which we can capture the sense of the internalised other is to say that it partly explains the development of our dispositions of character. By a disposition of character, I shall mean a relatively enduring disposition to respond reliably in a certain way when an appropriate situation presents itself. This means that when we say someone is of a compassionate disposition, we believe that she can be relied upon to have a particular set of thoughts and feelings when perceiving the suffering of another, and will be motivated to act to alleviate the suffering as a result of those thoughts and feelings. To say that the internalised other partly explains such dispositions means that the explanation of why some people come to be compassionate and others don't is dependent upon the ways in which our emotional development has been shaped by those around us. The compassionate person is someone whose emotions have been educated such that she perceives suffering in a particular light, and this education has been such that the disposition is reliable and enduring.[4]

With this in place, let me now return to the concept of shame. This emotion differs from certain others in that it there is widespread agreement among empirical psychologists that it is first seen at a

---

[3] See pages 56-62 above.

[4] For a brief account of the concept of dispositions of character, see Peter Goldie's *On Personality*.

later stage than emotions such as fear or disgust.[5] Opinions vary as to why this might be the case, one possibility being that shame requires cognitive capacities which develop later than those needed for what are often labelled 'basic emotions', and another being that shame requires a sense of self, which is only acquired during the course of the second year. I shall not be concerned here with trying to resolve this particular issue, but instead shall be focussing on how it is that we come to experience shame.

Allan N Schore[6] argues that shame begins to emerge when an infant behaves in a particular way, and looks towards the carer for the positive, affective response which she is accustomed to receiving. Instead, the look she receives is one of disgust, which the parents are apparently often unaware of themselves. This means that the positive, joyful expectation of the child is met with the 'other's unexpected refusal to cocreate an attachment bond that allows for the dyadic regulation of the emotion'.[7] Even at this stage, we see the development of an emotion which crucially involves the disappointment with regard to expectation, apprehended visually in the look of an important other. Schore argues that by the time we reach the end of late infancy, 'the elicitation of this affect does not require the presence of an external person', claiming that we have 'internalised...the eye of the self gazing inward'.[8]

The model which I have already proposed suggests that the input from the carer, which eventually comes to be internalised, will be systematic not only in terms of the manner of responding, but also in terms of the forms of behaviour which elicit the response. The same activities will be met with the same look of disgust or disdain time and again, resulting in our feelings towards certain states and activities being shaped in accordance with the response of an important other. In shame, we sense the look of disgust or disdain initially in the face of the carer, and subsequently in the feelings we have which would have been those of that person. Of course it would be ludicrous to suggest we always imagine another person seeing the activities of which we are ashamed, but as Bernard Williams points out, the internalised other 'is potentially somebody rather than nobody'.[9] This follows empirically from the observation that our feelings have been shaped by somebody rather than nobody, and it is therefore the responses of particular individuals who will influence what we come to see as shameful.

---

[5] For discussion of this see Lewis *et al*, 'Self Development and Self-conscious Emotions', reprinted in Jenkins, Oatley and Stein eds, *Human Emotions: A Reader*, 158-68, and Allan N Schore, 'Early Shame Experiences and Infant Brain Development', in Paul Gilbert and Bernice Andrews eds, *Shame: Interpersonal Behaviour; Psychopathology and Culture*, 57-78.

[6] See *op cit*.

[7] *ibid* 65.

[8] *ibid* 68.

[9] Bernard Williams, *Shame and Necessity*, 84.

There are other elements of the psychological material which are also instructive in terms of understanding the development of shame. Firstly, the idea of disappointment. The approval or support of the other is something we come to expect, and refusal of this can be crushing. One of the standard responses in shame is to look downwards – away from the gaze where we had hoped to find reward, but instead find disgust. Once this gaze has been internalised then we are no longer disappointing someone else, but disappointing an aspect of the self. Both Bernard Williams and Gabrielle Taylor talk of how shame induces a sense of powerlessness, which can be seen as stemming from the inability to escape the 'inward gaze'. In falling short of a standard which we wish to achieve, we are less than we want to be, and once this inward gaze has become part of who we are, we are powerless both in the sense of being unable to escape it, and in the sense of feeling less than the person we wish to be. The gaze that we have internalised conflicts with what we want, but in such a way that it not only clashes with, but judges the desires in question. To fall short of the standard is not to disagree or dispute it, it is to fail – to be worse.

This raises the further important point that the look of disgust which eventually comes to be internalised is encountered in opposition to our personal desires. The infant who meets with this look in his mother does so whilst enjoying a particular activity, and in the hope and expectation of reward. The response she receives *conflicts* with what is enjoyed and what is expected. Once again, the structure of these responses in the earliest stages mirrors what we see later on. The Spartan man who feels shame at being a coward experiences a conflict between his personal desire not to fight, and his attitude towards what is expected of him as a Spartan. In such cases we see not only how shame reflects a conflict between personal desires and the internalised other, but also how the latter can come to embody what are not simply the expectations of the primary carer, but what we tend to think of as shared ideals. It is to this aspect of shame which I shall now turn.

Shame and Shared Ideals

I have claimed that the views of those around us, principally the primary carer will influence the development of our emotional lives, which will include shaping which states and activities we come to see as shameful. This means that if the carer's attitudes reflect those of the wider community, we not only come to feel shame at activities which are disapproved of by those who bring us up, but also to participate in a wider 'honour code'.[10] To be honourable will be to meet the standards which are expected of us, and which are now a part of our overall perspective in the form of the internalised other. To be shameful will be to fall short of those standards, and we may experience this either publicly or privately. Although the standards may have been external to us in the first

[10] The use of this term is not intended to imply a rigid or codified set of rules, but is used in a broader sense of a set of norms indicating what is to count as honourable and shameful.

instance, once we have internalised them we no longer need a public in order to feel judged according to them.

This means that once we have acquired a sense of shame which reflects the attitudes of those around us in the wider community, we are bound to each other in a deeply personalised way. We expect certain forms of behaviour from one another, and are united in our opposition to other forms of behaviour. We judge ourselves and one another according to standards which we each accept as members of a group who share not simply beliefs, but a deeper emotional commitment to certain states and activities. I feel not only disgust at myself or at you when one of us engages in a shameful act, but at anyone who behaves in such a way. It is partly by virtue of the fact that we share these sorts of commitments that we see ourselves as members of what we think of as a group, rather than a collection of disparate individuals. Should I find myself amongst people who have radically different views on what is shameful or honourable, then it will place a severe restriction on the extent to which we can commune with one another.

So shame binds us within a community, but does so in a way which is dependent upon the very concept of community. I have argued that shame develops when our personal desires are met with disgust or disdain in the face of an important other whose gaze we come to internalise. It follows that shame could not develop without that initially external gaze. Without this, there is no opposition to personal desires, except conflicting desires. We would only have preferences to be gauged according to what brings greatest satisfaction, without the possibility of an external standard against which to measure them. But is it not conceptually possible that someone should establish personal standards to which she wishes to adhere at all times, and then feel shame at times when she falls below them?

The answer to this will depend on how extreme the example is. If we are to focus on a case of someone who has had no contact with human society, then the answer must be no. In the first place, an absence of contact with any other human beings will bring into play standard Wittgensteinian concerns about how public language makes possible our ability to identify an inner state such as shame. Without the language which a community makes possible, even formulating thoughts about standards to be maintained becomes unimaginable. Adam Smith presents this point more prosaically.

> Were it possible that a human creature could grow up to manhood in some solitary place, without any communication with his own species, he could no more think of his own character, of the propriety or demerit of his own sentiments and conduct, of the beauty or deformity of his own mind, than of the beauty or deformity of his own face. All these are

> objects which he cannot easily see, which naturally he does not look at, and with regard to which he is provided with no mirror which can present them to his view. Bring him into society, and he is immediately provided with the mirror which he wanted before.[11]

I have suggested that shame must be seen as necessarily social, in that it is only once we engage with other persons that the structural requirements become available.[12] It is not clear to me that wherever there is a community there must be shame, in that one can imagine a possible world in which no-one ever showed disdain for anything another person did, or another in which disdain was so arbitrary that we could not come to see particular states or actions as shameful. Nevertheless, the widespread presence of shame across different cultures[13] suggests that it is a likely consequence of the human condition that where cultures emerge there is likely to be a sense of shame which develops with them.

With regard to the notion of shame binding us within a culture, Bernard Williams offers a similar argument in *Shame and Necessity*, and elaborates on it with reference to the fate of Ajax.[14] Infuriated after Achilles' arms are given to Odysseus instead of himself, Ajax decides to kill the heads of the army, but is rendered mad by Athene before he can do so. In his state of insanity, he slaughters sheep and cattle in the belief they are other Greek soldiers. On awaking the following morning his mind has been restored and he now has a vivid memory of what he has done. The realisation of what has befallen him inspires a sense of shame so powerful that he commits suicide. The language used in Sophocles' description of the event reflects the nature of shame as already discussed. When considering his plight, he says,

> What countenance can I show my father Telamon?
> How will he bear the sight of me
> If I come before him naked, without glory,[15]

It is the gaze of his father which torments him, with the image of his nakedness symbolic of the way in which his character is exposed in such a way that he is seen having lost his honour. But an understanding of why the shame is so intense, or even why there is any shame at all, requires an appreciation of the standards of honour which have become part of Ajax's understanding of who

---

[11] Adam Smith, *The Theory of Moral Sentiments*, in D D Raphael ed, *British Moralists 1650-1800*, vol I, §801.
[12] Although I have spoken in terms of communities and societies, there is no reason why the groups involved need by any greater than two on the account I am offering.
[13] For specific discussions of shame in different cultures see 'Gender, Shame and Culture: An Anthropological Perspective', by Nancy Lindisfarne, and 'The Sacred and the Social: Cultures of Honour and Violence', by Cohen, Vandella and Rantilla, both reprinted in Paul Gilbert and Bernice Andrews eds, *Shame: Interpersonal Behaviour, Psychopathology and Culture*.
[14] The following summary is derived from chapters three and four.
[15] Quoted from Bernard Williams, *Shame and Necessity*, 85.

he is. The internalised other requires him to live in a certain way, and to meet certain standards of conduct. In Williams' words, 'the interlocked expectations between himself and the world are of course peculiarly connected with his status as an heroic warrior, and that is why, in his case, his grotesquely unsuccessful and ridiculous attempt counts for so much.'[16] To have internalised the standards associated with being an heroic warrior involves the acquisition a particular way of interpreting one's own actions. Ajax's sense of honour and worth are intimately tied up with these standards, and to fall so far below them means he simply cannot but see himself as shamed and dishonoured.

What emerges from this example is the way in which public ideals come to partly constitute who we are through the internalised other. The Homeric world had certain ideas of what was required of a hero, and Ajax came to embody these ideals. With this process comes both a way in which he interpreted himself, but also a way in which the world saw him and understood him. As such, he was immersed in culture in which he both saw and was seen in a particular light, and where he both expected himself, and was expected by others, to achieve certain standards. The internalised other bound him to the world around him by tying his own way of being to the way in which others relate to the world which he shared with them.

I have focussed on Williams' discussion of Ajax's shame, but I suggest that the key arguments he advances there generalise. The development and structure of shame which put forward here offer a context in which we can understand the plight of Ajax, and the specific example is one which illuminates how not only the ideals of Greek nobility, but a variety of shared ideals can come to be part of the perspective of individuals who are raised within a community. Once certain ideals are present, they will constitute the standards we come to internalise, such that those standards become part of us. The way we interpret the world, and the way others interpret us are now bound up with one another. There are shared standards which establish the possibility of my feeling shame at being revealed as a coward, and others seeing me as shameful for my cowardice. The existence of these shared standards is the condition for the possibility of shame.

Now, these points concerning the development of shame reveal a number of issues of much wider significance. Firstly, they provide an account of how it is that the values of others can come to shape our own perspective on the world. From an early age, we begin responding to the behaviour of others in ways which lead to our coming to respond to the world in ways shaped by their influence over us. Aspects of this behaviour fall into what we intuitively think of as the moral domain, and we therefore have an indication of how we might argue that moral values are transmitted. What is now required is a broadening of the discussion, so that we expand on the

---

[16] *Ibid.* 85.

discussion of the development of the internalised other with regard to shame, and focus on its relevance to the idea of a wider moral perspective.

Moral Development and the Internalised Other

I have argued throughout that our emotions are developmentally open, to the extent that the objects to which we respond are open to influence by those around us, a phenomenon which is particularly important in explaining our emotional development at very early stages in our lives. The way in which this operates is that facial expressions, tone of voice and other forms of behaviour are such that we grasp the attitude of the other. But consider once again the term used by Davidson when explaining how it is we come to learn the names of objects. When we get something right, we receive a 'reward'. Reward is not something that can be reduced to an articulation of 'that is correct' or 'that is accurate'. To be rewarded is not simply to be acknowledged as being correct or accurate, it is be acknowledged as having done something which is *better*. The response we receive is a positive one, and as such it expresses a form of endorsement. Similarly, a wide range of other reactions which we grasp have evaluative dimensions buried within them. If a parent shows anger, we grasp not simply displeasure but also a sense that this activity is *worse*. Many standard responses work in this way, shaping our feelings towards certain states and activities as a result of our grasping that they elicit reward or punishment (which I shall use in the same metaphorical sense in which Davidson uses the concept of reward) from those whose responses have no importance for us.

In a large class of cases we are rewarded or punished for activities which we come to understand are of benefit to us personally. The horror or anger at a tendency to flirt with danger is something we come to recognise as a response guided by questions of what is in our own self-interest. We come to internalise the fear of certain objects in light of the responses from those around us, but this is a fear which relates to our own interests.

Many of the rewards and punishments we receive relate to a domain of interest which can only be understood as standing in contradistinction to self-interest. We are rewarded for activities which we cannot understand as being necessary for us to flourish at a purely personal level, and we are punished for actions which do not harm us in any way. This means our background of directed feeling comes to orient us towards a wider world than that of self-interest. We reach a stage where we see fire as dangerous, not because of a recurring thought about the properties of fire on each occasion, but through habit as a result of receiving the same response time and again when we approach it.[1] Similarly, there are certain activities and states for which we are consistently punished

[1] Needless to say, personal experience can also shape our feelings, in that the child who has been burnt by fire is at least as likely to 'see' it as dangerous as the one who hasn't.

or rewarded, which also shape the way we see the world, but which are not comprehensible in terms of what benefits us in ways reducible to self-interest. It is this second domain which I characterised as the 'internalised other', and it is within this domain that we find the ethical.[18]

On this account, an ethical outlook is to be understood as an orientation towards the world which stands in contrast to that determined by self-interest. It is when our aggressive actions meet with a look of disgust and anger, and our compassionate ones are met with encouragement and enthusiasm that we come to see aggression, suffering and compassion in a particular way. We are no longer simply responding to the suffering of another on a purely personal impulse, we act from a disposition which reflects the influence of those around us. A response of compassion is now not simply a personal response to suffering but one which finds favour with the internalised other, one which reflects the way we have learned to feel about the world. It is this second dimension which brings it within the domain of the moral. Only the feelings we have internalised make possible the relevant new beliefs about such actions, in that they may now be perceived from a perspective which has been broadened by the presence of the internalised other. From the wider perspective of how others see us, which has now come to be part of our own, there is an alternative to our personal desires. Running away is no longer just saving myself, but letting others down. Self-preservation now has a rival in the form of wider expectations of how to behave, and it is from within this wider perspective that fleeing can be an action which we see as cowardly.

In short, the moral perspective involves acquiring as a part of our personal identity a recognition of the claims of others. This is a function of our emotional development, and it means that we come to feel a certain way towards others, ensuring that their interests weigh in the balance when we come to consider the world around us. These claims stand in contrast to those of personal self-interest, and they come to be a part of us as a result of constant interaction with others around us and the impact this has on our emotional development. It follows that insofar as we could imagine someone whose emotions have not been shaped by those around her, this person could not have a moral perspective. Given the claims in chapter one concerning the irreducibility of emotions to cognitive states, it also follows that all cognitivist theories of ethics are wrong insofar as they take the moral perspective to be one which involves non-feeling laden beliefs about the world.

The characterisation of the moral perspective which I have offered is that it is the recognition of the claims of others which are part of our personal identity, and which can be understood in contradistinction to the goals of personal self-interest. This perspective is acquired as a part of our emotional development and partly-determines the way we perceive the world around us. On this

[18] I say the ethical is 'within' the domain of the internalised other rather than 'equivalent to' as one might wish to argue that the source of aesthetic judgement is also located here.

account, there could be a wide range of different sorts of content which would satisfy the criteria for a moral perspective as I have defined it. Christian asceticism, Greek heroic values, liberal individualism and a range of other such sets of values could all provide the basis for inculcating in us a recognition of the claims of the wider community, and in each case the person who expresses views in line with these values has a moral perspective. This also indicates how and why there is the possibility of moral conflict and moral dilemmas. If the content of morality is contingent in the way suggested here then it is to be expected that there will be different moral perspectives both between and across communities. There is also an indication of how inner conflict could occur. Most obviously, if personal desire and the internalised other are defined in contradistinction to one another, then there is possibility of an opposition between what we want and what we ought to do - between personal desires and the claims of others which are now part of our own identity. It is this sort of conflict which is captured in terms of the debate over selfish and selfless behaviour, and the account offered here indicates this is probably an inescapable part of our moral development. Even the person in whom morality always wins out is likely to feel the pull personal desire on some occasions when such desires conflict with their view of what they ought to do.

A further potential for inner conflict emerges given the dependence on wider cultural influences in the development of the moral perspective as I have characterised it here. There just may not be a fully consistent set of claims which come to shape our identity. The general complexities of any human culture mean that we may find ourselves reflecting sets of values which lack a fully internal consistency, or else which fail to correspond to all the facts we find before us. In such circumstances, the need for some sort of adjustment in the content of our moral views will press in on us. In both the sorts of cases set out above we see the potential for disagreement with the internalised other, and through this either a threat to our moral outlook, or a need to adapt it. What we need to look at now is whether or not we can demonstrate whether or not there is a certain content our moral outlook ought to have, and thereby move towards at least the possibility of greater internal consistency, and consistency with the world around us.

5. Moral Content

Introduction

The previous chapter provided a formal description of what is it to have a moral perspective, one which relies upon the metaphor of the internalised other, and which demonstrates how it is that our moral development is dependent upon our emotional development. The aim of this chapter will be to extend the discussion into the area of moral content, and more specifically to try to determine the kind of content the internalised other ought to have - to specify the content of the judgements which we ought to form from the moral perspective. The problem with such an endeavour is in trying to justify claims along the lines that being kind or sympathetic are examples of goodness, without relying at least implicitly on a moral theory which already has such claims among its premises. The danger of making claims such as 'we ought to be kind because it is good to be kind', is that we can simply be accused of begging the question, and we therefore have a reason to offer a further line of argument to defend this sort of claim. In order to highlight the nature of this problem, I shall begin by looking at the arguments set out by Schopenhauer in defence of the claim that compassion is the basis of morality. I shall suggest that his claims, similar in many respects to those set out by Hume in his defence of an ethics based on the sentiment of benevolence, are vulnerable as they stand to a series of telling counter-arguments.

The aim of the second half of this chapter will be to supplement the sentimentalist line in such a way that it can overcome these responses, and this will lay the groundwork for the claim that compassion and reason are the basis of a proper morality. The way in which I shall try to justify this position is by appealing to a naturalistic psychology. I shall show that once one takes account of certain features of the way humans are - our need for others and our capacity to respond to others' needs, then a morality based around compassion is one that makes sense in the light of such considerations. This will provide the further line of defence against the counter arguments to sentimentalism which I shall suggest pose a considerable threat to such a moral theory. To bring this out, let us consider one classic example of a sentimentalist, ethical position.

Compassion and Morality

One way of addressing this issue is to ask whether or not compassion could account for the whole of morality. This is the position taken up by Schopenhauer in his argument for an alternative to Kantian moral philosophy.[1] He argues that 'It is simply and solely this compassion that is the real basis of all *voluntary* and *genuine* loving-kindness. Only insofar as an action has sprung from

compassion does it have moral value; and every action resulting from other motives has none.'[2] The two italicised words are to emphasise that we are only considering actions which we intend, as against those which fortuitously result in the alleviation of someone's suffering, and that we are only interested in those actions where concern for someone else's 'weal or woe'[3] dominates, as against a desire to appear concerned by such issues. He identifies three sources which account for our entire range of motivations. The first is egoism, which is defined as a desire for one's own well-being. The second is malice, which is the desire for the pain of another, but without any benefit for ourselves. And the third is compassion, which is the desire for another's well-being.

It is claimed that compassion is aroused by the suffering of others, and relates to others' happiness only because 'the nature of satisfaction, enjoyment, and happiness consists solely in the removal of a privation, the stilling of a pain; and so these have a negative effect.'[4] The reason that we are not moved by the 'fortunate and contented man' is that the absence of 'pain, want and distress' leaves nothing to arouse our compassion. Compassion therefore works in two ways, motivating me to help others who are in difficulty, and preventing my bringing about suffering. With these two foundation stones in place, Schopenhauer is able to set out what he considers to be the 'supreme principle of ethics', which is '*Neminem laede; imo omnes, quantum potes, juva.*'[5] The sight of someone else suffering can move me to help them, but the thought of my causing suffering is also sufficient to counteract malicious or egoistic drives, and to prevent me from inflicting suffering on others. Where compassion prevents me from injuring others, we see the basis of justice and the duties of law which stem from it. Where compassion moves me to help others we see the basis of philanthropy, and the duties of virtue which stem from it. Both parts of this twofold division can be traced back to compassion.

Not only justice and virtue, but also right and wrong are defined with regard to compassion – 'the concepts of *wrong* and *right* are synonymous with doing harm and not doing harm, and to the latter belongs also the warding off of injury'. When we talk of evil, it is to be understood in terms of causing greater harm, with the greatest evil being those actions which deliberately cause the greatest degree of suffering.

These are little more than the bare bones of a rich and powerfully argued piece, but the way in which one might construct a wider theory of morality on the foundations of compassion is set out in a way which often achieves immense intuitive appeal. We have a 'natural compassion' which

---

[1] Arthur Schopenhauer, *On the Basis of Morality*.
[2] *Ibid*, 144
[3] Where the characterisation of compassion set out earlier differs from Schopenhauer's is that I do not think that 'weal' is necessary here.
[4] *ibid*, 146.

means we are moved by the suffering of others, and moved to avoid causing suffering. We are provided with an account of the psychology of the virtuous agent as someone who is disposed to help others and to avoid harming them, and with an explanation of how other moral concepts such as duty and justice emerge out of compassionate natures. The fact that compassion is present in all of us also explains the widespread presence of morality. There is one further point here which will be of central importance to the defence which I shall be mounting later for what is, broadly speaking, the arguments set out by Schopenhauer. Suffering reveals need, and compassion is a response to need. Compassion would be less important if we were always or even usually capable of ending our own suffering, but this is often not the case. In many of the cases when we suffer, we do so in ways which are typical for any human being - physical pain, grief, disappointment, loss - these are all typical aspects of what we must endure, and a compassionate response is therefore one which is called forth by the kind of creatures we are, and by the kind of needs we typically have.

To what extent, then, can we accept Schopenhauer's claim to the effect that compassion is the basis of our ethical life? One way of addressing this is to pose the Socratic question with which Bernard Williams opens *Ethics and the Limits of Philosophy*. If one is to answer the question 'how should one live?', could it be answered entirely in terms which are rooted in compassion? There are many cases when this is certainly true. If we take helping others who are in distress, or not causing distress ourselves as important aspects of a moral outlook, then our concern about the suffering of others can provide a plausible explanation of why we hold these maxims to be true. But are there not cases where we behave in a moral or immoral fashion which involve no harm or suffering? If we take the case of a man who feels guilt because he has failed to visit his mother's grave for a number of years, or a woman who is ashamed of herself for being too lazy to attend the church fundraiser, are these not perceptions of immoral behaviour on the part of the agent which might involve neither harm nor suffering? One could elaborate on the stories in such a way that the man imagines how his mother would have responded were she alive, or the woman imagines the disappointment of the priest, but such addenda are certainly not necessary. Each of them could say quite coherently that they firmly believe they have behaved immorally, but no harm has been done.

One could argue that any failure to meet a moral standard is in some sense painful for other members of the community, or harmful to oneself. But this is partly circular and partly wrong. If morality is ultimately to be defined in terms of causing and alleviating suffering, then we cannot define instances of suffering in terms of moral failure. In any case, even though seeing someone else's immoral act may pain us, it is surely not wrong because it pains us - it pains us because it is wrong.

---

[5] 'Injure no one; on the contrary, help everyone as much as you can', *ibid* 147.

The suggestion here is that there are aspects of morality which seem to stray outside the bounds of harm or suffering, thereby raising serious questions as to whether compassion can provide a full account of morality. In order to highlight this point, consider these two extracts from Kitto's description of life in Sparta.

> The Spartiate was forbidden to engage in agriculture, trade or professional work: he must be a professional soldier. He had his farm, worked for him by helots, he dined in public 'messes', to which he contributed his share from his farm: if he failed to contribute, he ceased for the time being to be a full citizen. Family life was severely limited. Babies adjudged weak were done away with; boys lived with their mother until they were seven; from seven to thirty they received the appropriate kind of public military instruction and exercise. Girls too were given careful physical training.[6]
>
> Sparta was admired for her *Eunomia*, her 'state of being well-lawed', because – whether you liked her ideal or not – she did through her laws and institutions train her citizens in this ideal with unusual completeness: she did train citizens selflessly devoted to the common good:'[7]

The problem with such a picture for anyone who wishes to explain morality in terms of compassion is that we are presented with a Spartan culture which has a notoriously rigid moral code, but one in which it scarcely looks as if compassion is doing a great deal of work in propping it up. If one were to single out the fundamental concepts which capture Spartan morality, then those which spring to mind would surely be ones such as honour, duty, heroism, and other virtues which contributed to the warrior code which symbolised Spartan culture. If one accepts that such concepts contributed to the normative life of ordinary Spartans then there is a problem for the claim that compassion grounds morality, in that these concepts just don't look as if they are reducible to compassion. One can certainly acknowledge that there may have been more compassion in this community than the standard historical accounts suggest, but the picture of a fearsome, warlike minority with an all-embracing sense of the common good involving hideously tough military and other physical training simply doesn't look as if its morality could be understood exclusively in such terms.

There is a further serious problem for any such thesis when one comes to consider the meaning of moral concepts. Schopenhauer tells us that right and wrong are synonymous with doing harm and not doing harm, and as compassion is the basis of all morality, then one can infer that all other moral concepts will also be analysable in terms of harm and suffering. More specifically, good will

---

[6] H D F Kitto, *The Greeks*, 91.
[7] *Ibid*, 94.

be analysed in terms of alleviating and not causing suffering, and bad in terms of causing or failing to alleviate it. But if this is so, then propositions such as 'suffering is good' would constitute a logical contradiction, and this just doesn't seem to be the case.

When Nietzsche rails against the 'levellers' on the grounds that they 'take suffering itself as something which must be *eliminated*', and says that 'harshness, violence, enslavement.....heighten the species human being'[8], we may recoil, but one can nevertheless read into what he says a coherent moral philosophy. In a later section of the same text, we are told that 'all this violence, arbitrariness, harshness, horror, nonsense has turned out to be the means by which the European spirit was bred to be strong, ruthlessly curious, and beautifully nimble.'[9] Because there have been protracted periods of obedience to despotic laws, 'something emerged and emerges that makes life on earth worth living: virtue, for example, or art, music, dance, reason.'[10] In direct opposition to Schopenhauer, Nietzsche can be read as suggesting that suffering constitutes a necessary means to making life worthwhile, and must therefore be embraced rather than eliminated. My point here is not to defend Nietzsche, but merely to highlight that if one accepts Nietzsche has provided an alternative moral philosophy which gives suffering and harm important positive roles, it just cannot be true *a priori* that suffering is bad.

There is one further step which can be taken down this path. One interpretation of Nietzsche could be that those of a 'higher nature' care nothing for the weak, and thrive on uninhibited expressions of the will-to-power through media such as art or warfare. Compassion is a restriction on the expression of the will-to-power and is therefore to be set aside in order to allow the unfettered release of this force. On such a reading, compassion could be presented as neither necessary nor sufficient for morality, but more than this, it could even be construed as a vice.

The problem emerges when we come to consider a statement such as 'compassion is the basis of morality', which anyone would have to defend as true in taking something like Schopenhauer's line. If one defends this as an *a posteriori* claim, then it is open to counter-examples such as the Spartan one - it just isn't the case that all instances of morality can be understood in terms of compassion. If one defends it as an *a priori* claim, then one can argue that the Spartans just didn't have a full grasp of the concept of morality and failed therefore to give compassion the place it deserves in their moral system. This cognitive failure explains why their morality appears twisted to those more attracted to Schopenhauer's position. But this merely invites the next problem, which is that if 'compassion is the basis of morality' is true *a priori*, then 'compassion is not the basis of morality'

---

[8] Friedrich Nietzsche, *Beyond Good and Evil*, § 44.
[9] *Ibid*, § 188.
[10] *Ibid*

would be necessarily false, and this just doesn't seem to be the case - there is no *logical* contradiction involved here. We may recoil at the thought of a moral system which ignores the suffering of others, but we could still recognise it as logically coherent example of a moral system. A compromise between the two positions along the lines that compassion plays a major role in the moral system of many cultures appears easier to defend as an empirical observation, but we now encounter even more problems. We now need an account of what explains those features of moral systems which cannot be explained in terms of compassion, and we need an account of how compassion does relate to morality given that it is no longer being defended as the basis of it.

If these arguments are correct, any attempt to construct an *a priori* argument to the effect that all moral concepts derive from the meaning of compassion still needs a lot of work, and *a posteriori* arguments to the effect that where there is morality there is compassion at the heart of it just appear to be false. In light of this, we need to be clearer about how compassion comes to be related to morality, and we must reject that claim that merely by showing that compassion is naturally occurring in human beings we have provided an account of how morality is possible and what it is. The relationship between the two is just far more complex than such an account suggests.

A Proper Morality

The previous chapter characterised the moral perspective as a feature of the internalised other - the part of our identity which is shaped by the influence of others. My claim was that we have the capacity to have our feelings towards certain states and activities shaped by those around us, and the orientation this provides us with, where it is in contrast to self-interest, includes the domain of the moral. So far, this is a purely formal description of what is meant by having a moral perspective, and how it is that this perspective is acquired. But we now need to give an indication that there is a certain sort of content that the moral perspective ought to have. The difficulty here is one of trying to justify a certain sort of moral perspective without recourse to precisely the sort of moral values one is trying to justify. The strategy I shall adopt is that a moral perspective which is based in compassion can be defended from the standpoint of a naturalistic psychology. Given the way humans are, then compassion is the proper basis for our morality.

In the first instance then, can we show how the content of the internalised other can be constrained in some way so that what we think of as paradigmatically immoral views can be excluded? An initial response is that one explanation of the hideous atrocities which we have witnessed in human history is that the perpetrators were acting to achieve a moral goal. The theory provides an explanation of how such moralities, appalling as they may strike us, come to shape the views of large numbers of people and to provide them with the motivation to carry out the actions in

question. But surely we want a moral theory to do more than this, in that we want not simply a description of the psychological processes involved and formal conditions, but also a normative dimension - we want to see how such positions which are described as morally right by some can be shown to be morally wrong. We want to be able to argue that there is something at least approaching a proper morality whose content we can go some way towards specifying.

One limiting factor is the issue of what is true independently of our moral views. I have argued that emotions are partly-constituted by beliefs or perceptions, and when the feeling is combined with a belief which is false or an inaccurate perception, we judge the emotion to be inappropriate. In a case of fear, if we wrongly believe the snake slithering towards us is poisonous, then we have good cause to think fear is no longer the proper response to the situation.[11]. This much is consistent with Hume's treatment of the passions.

> |I|t must follow, that passions can be contrary to reason only so far as they are *accompany'd* with some judgement or opinion. According to this principle, which is so obvious and natural, 'tis only in two senses, that any affection can be call'd unreasonable. First, when a passion, such as hope or fear, grief or joy, despair or security, is founded on the supposition of the existence of objects, which really do not exist. Secondly, when in exerting any passion in action, we chuse means insufficient for the design'd end, and deceive ourselves in our judgement of causes and effects.[12]

Many of the moral ideals we wish to rule out may well be vulnerable to precisely these sorts of attack. If a moral outlook entailed some pseudo-scientific theory of the superiority of certain racial groups, it would be relatively straightforward to point out the factual mistakes on which the attitudes depended, with a view to forcing a change of heart. But even if such a strategy were to have some effect, we still need more. If, as I have suggested, the moral perspective is to be understood as an orientation towards the world dependent on certain representations, then we want ways of ruling out this sort of *orientation*, as against the means on which it depends for its sustenance. In other words, we want to be able to say that anti-Semitism is wrong as a type of orientation, as against anti-Semitism is wrong because Jews do not in fact have an unusually shaped cranium.

---

[11] Perhaps some qualification is needed here, in that there may be complex cases where we have the an appropriate emotion despite having a relevant false belief, and an inappropriate emotion despite having a relevant true belief. If I fear the snake's bite, and the snake is actually harmless but happens to belong to someone who shoots all those who ever catch sight of this treasured pet, then there is a case for arguing that I am experiencing an appropriate emotion despite holding a false belief. Conversely, if I am overjoyed on recognising the snake is harmless then one could argue I am feeling the wrong emotion based on a relevant true belief. Such cases will tend to emerge when the emotions are partly consituted by beliefs or perceptions which reflect only part of a wider picture, and an absence of the full range of relevant facts means an agent can have emotions which are appropriate given the information available to the agent, but would be seen as inappropriate if all the relevant facts were known to her.

In order to achieve this, we need to identify a means of specifying how the structure of the internalised other can itself justify the exclusion of certain attitudes, and thereby regulate the content of moral theory in a way which does justice to the intuition of the previous paragraph. One starting point can be found in the work of David Wiggins.[13] I have suggested we cannot identify our moral responses with certain emotional responses to independently existing properties in the external world, but we can go some way in explaining why certain properties are likely to come to elicit the responses which are part of the moral outlook we have within the parameters of the overall thesis I have set out.

We can begin with a question which Wiggins asks. 'Is it not rather there is something in the object that is *made for* the sentiment it would occasion in a qualified judge, and it brings down the sentiment upon the object as so qualified?'[14] In order to highlight this, consider the following scenario. On impulse, a child lashes out and strikes his mother in the face. The mother darts back a look of anger, and scolds the child. The same thing happens again with the same response on the part of the mother. The child grasps the displeasure in the mother's response and associates it with his violent action. Over time, as similar actions receive similar responses, he comes to develop certain feelings towards violent actions which guide his behaviour in the form of inhibiting such actions in himself and disapproving of them when committed by others.

In the context of the thesis set out here, the feelings which the child has developed are part of the internalised other. Such an account aims to show how the response of another shapes and broadens our perception of ourselves and those around us. This much has already been argued for, but there is a further aspect to the scenario which is also of importance. The reaction of the mother is one with which most of us could identify, and this is the case because anger at being struck is a response which makes sense *given the kind of beings we are*.[15] This means that reactions and behaviour which determine the content of the internalised other will bear the marks of our humanity. Wiggins asks us to suppose that

> [o]bjects that regularly please us or help amuse us...or harm or annoy of vex us....in various ways come to be grouped together by us under various categories or classifications to which we give various avowedly anthropocentric names; and suppose they come to be grouped

---

[12] David Hume, *A Treatise of Human Nature*, 416.

[13] David Wiggins, 'A Sensible Subjectivism', in his *Needs, Truth and Values*.

[14] *Ibid*, 193-194

[15] The idea of 'what makes sense to us' is central to the arguments of Alan Gibbard in the first part of his *Wise Choices, Apt Feelings*.

> together *because* they are such as to please, help, amuse us, ...or harm, annoy, vex us...in their various ways.[16]

In the context of the ethical, the impact of this argument would be that there are certain properties of the external world and the human organism which are likely to produce what Wiggins calls '<property,response>' pairs in us. Anger at being struck, joy at being helped, suffering at seeing the pain of someone close to us. It is these marks of the kind of creatures we are which will condition our own responses, and through those responses, shape our own feelings and those of people around us. Once we have come to internalise the feelings of opposition to causing pain, then we are liable to respond in a particular way once we are aware of having caused it. Guilt becomes the standard response for the agent whose has come to perceive violence in a way which is guided by a background of directed feeling, which has in turn been shaped during scenarios such as the one set out above. What we perceive is now laden with the feelings which are directed towards such specific actions, such that perception and sentiment are phenomenologically united.

Much of the account I have offered is consistent with what we find in Hume, but one difficulty posed by Hume's account is the famous distinction he draws between reason and taste. Reason discovers objects 'as they really stand in nature', whilst taste 'has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises, in a manner, a new creation'.[17] This would be a major hurdle in the quest to establish that ethical propositions are true descriptions of the world, and that values are genuine properties of the world. But this is not my aim here. I wish to establish that there are properties in the world which elicit what we can judge as appropriate responses, and that these responses in turn condition the content of the internalised other. In other words, there is a certain content which our ethical outlook ought to have given the way we are, where the ought in question is not one of morality, but one springing from a naturalistic psychology. There is a certain way one would expect morality to be given what human beings are like.

Far from raising difficulties in such a quest, Hume provides numerous examples of exactly how this might work. When it comes to judging questions of taste or beauty, we 'feel a sentiment of complacency or disgust, according to the nature of the object, and disposition of our organs'.[18] Similarly, '*cleanliness* is also to be regarded as a virtue; since it naturally renders us agreeable to others'.[19] This last example is at odds with the moral theory I have presented in that I have suggested that such properties only come into the moral realm when judged virtuous by the

---

[16] David Wiggins, A Sensible Subjectivism, in *Needs, Truth and Values*,195.
[17] David Hume, *An Enquiry Concerning the Principles of Morals*, 88.
[18] *Ibid*, 86.
[19] *A Treatise of Human Nature*, 611.

internalised other, but Hume's point is consistent with my claim in this instance. Simply by virtue of being the way we are, there are certain properties in the world which elicit specific responses. We prefer it when people are hygienic because it is more pleasant, and when they aren't we are likely to react with disapproval. This in turn is grasped by the perpetrator who comes to see personal hygiene in a particular way as a result of the feelings which now guide her perception of it.

The Exclusion of the Immoralist

At this stage it requires further consideration of how we might supplement the account of morality given by Hume or Schopenhauer in favour of what I shall call for the sake of brevity, conventional morality. By this I shall mean the view that the moral person is one who is concerned by the needs and suffering of others, and the moral act is one taken out of this sort of concern. Is it not the case that the insertion of the internalised other into the debate is simply a circuitous route to exactly the same problem? I have suggested that the conventional moralist is vulnerable if she relies on the claim that compassion or benevolence are good, where this is taken to be an *a priori* claim, and one possible reading of Hume is that he does more or less just this. There are points when he argues that morality is to be understood broadly in the way I have just described. It is part of our make-up that we approve of the qualities which increase happiness, and disapprove of those which promote misery. Our morality is a reflection of this tendency, and we can see this is in our moral terms.

> General language, therefore, being formed for general use, must be moulded on some more general views, and must affix the epithets of praise or blame, in conformity to the sentiments, which arise from the general interests of the community.....and even in persons the most depraved and selfish.....must attach the notion of good to beneficent conduct, and of evil to the contrary.[20]

It is hazardous to attribute to Hume a semantic analysis of our moral concepts, but one possible reading of the above passage combined with the one below from the same page of the *Enquiry* suggests that he believed it to be true *a priori* that, *ceteris paribus*, what promotes welfare is good.

> If we consider the principles of the human make, such that they appear to daily experience and observation, we must, *a priori*, conclude it is impossible for such a creature as man to be totally indifferent to the well or ill-being of his fellow-creatures, and not readily of himself, to pronounce, where nothing gives him any particular byass, that what promotes their

---

[20] David Hume, *An Enquiry Concerning the Principles of Morals*, 49

> happiness is good, and what tends to their misery is evil, without any further regard or consideration.[21]

The picture we are offered is of a human being with the sentiment of benevolence as part of a 'natural constitution', which leads us to value the qualities in people which increase happiness, and dislike those which promote misery. Our moral concepts reflect this underlying structure of thought and feeling. The meaning of good is determined by the sentiment of benevolence, and

> Had nature made no such distinction, founded on the original constitution of the mind, the words *honourable* and *shameful*, *lovely* and *odious*, *noble* and *despicable*, had never had any place in language[22]

This represents an ideal solution to the relationship between ethics and the emotions, in that if it were the case that 'pain is bad, and compassion is a response to pain' were true *a priori*, then we would have a necessarily true statement about a 'natural' emotional response to certain features of the world, with this response being part of our moral lives. This is the position I take Hume to hold. It is also the position which one finds more subtly in one of the recent major works of a sentimentalist tendency. Simon Blackburn asks us in the midst of his attack on intuitionism, 'Why not instead just see there being people who are motivated to care about the promotion of happiness, and who as a result are motivated towards the act, and, *equivalently*, are inclined to see it as good?'[23] (italics added).

There are two paths a sentimentalist can take at this point. The first is that we *can* analyse our moral language, and this will reveal a series of true propositions such as 'what promotes welfare is good'. The second is that human beings are constituted in such a way that they respond positively to behaviour which promotes welfare, and negatively to behaviour which promotes misery.[24] Our attitudes, as determined by these sentiments, constitute our moral lives, and these attitudes *cannot* be analysed.

It is in the move from the concern about the welfare of others to the concept of good where the problem lies on either interpretation. If we have a natural tendency to care about the plight of others, and this determines the content of our moral concepts, then it ought to be the case that we

---

[21] *Ibid*, 49.

[22] *Ibid*, 39.

[23] Simon Blackburn, *Ruling Passions*, 90.

[24] Such a position is more difficult to reconcile with Hume, who discusses cases such as wishing misery upon our enemies. Nevertheless, the *ceteris paribus* clause from the quote on the previous page could be interpreted as a means of explaining just such instances, thereby allowing us to qualify the broader claim that whatever promotes welfare is good, and whatever promotes misery is bad, whilst retaining the general character of the theory.

just have to think compassion is good. On the first path, this will be a matter of logical necessity given the meanings of the terms, and on the second path it will be a matter of causal necessity given the way the mind works. But neither seems to be the case, in that there is no logical contradiction in the statement 'promoting human welfare is bad', nor does it seem to be the case that everyone we consider to be sound of mind responds sympathetically to the misery of others.

The first path is one which I considered in response to Schopenhauer, where I argued that the lack of any logical contradiction in the claim that compassion or sympathy are vices, or that suffering is good indicates that it cannot be the case that statements such as 'compassion is a virtue' are true *a priori*. On the second path, our morality just is our tendency to wish to relieve the suffering of others and promote their well-being, and we go wrong when we look beyond this straightforward aspect of the human condition. Simon Blackburn presents us with such a position.

> We start off with something bad such as the piano being on your foot. An agent is concerned about that. A novelist might leave it there: we have a piece of sympathy and a right attitude, a modest indication of the agent's ethics. But the philosopher worries that this is very *mere* attitude or emotion on the part of the agent.[25]

Blackburn appeals to us not to look beyond certain attitudes to other persons for the explanation of our morality, as it is the further attempts to look for a truth-apt ethical proposition which creates much of the unnecessary confusion which afflicts moral philosophy. Much of what I shall say is sympathetic to his expressivist project, but this position strikes me as vulnerable as it currently stands.

If morality is to be understood as partly constituted by responding sympathetically to the pain of others, then how are we to answer the claim that morality is partly constituted by an unsympathetic attitude to the pain of others. When, for example, Nietzsche called for an end of sympathy for the 'surplus of deformed, sick, degenerating frail, necessarily suffering individuals'[26] whom he believed had come to dominate European culture, how can we refute his claims to have identified a superior morality to that of Schopenhauer or a 'flathead' like Hume? If Nietzsche and the conventional moralist agree on what compassion is, and on the fact that it has moral status, but then diverge on whether it ought to be deemed moral or immoral, and if all we have are attitudes which lie beyond analysis, we have no way of breaking the deadlock in favour of conventional morality. We need a further argument to get us past the point where we are simply sat opposite the immoralist thumping our fists on the table as we insist that we are right, and unless the account I am offering helps to

---

[25] Simon Blackburn, *Ruling Passions*, 91.
[26] Friedrich Nietzsche, *Beyond Good and Evil*, § 62.

exclude the immoralist then it will be open to the charge that it fails the normative test we would expect any moral theory to pass.

We can start by recalling Hume's observation that 'it is impossible for a creature such as man to be totally indifferent to the well-being of his fellow creatures'[27] This view is echoed in those thinkers whose view of human nature is perhaps most pessimistic. Even Hobbes will allow that '(O)f things held in propriety, those that are dearest to a man are his own life, & limbs; and in the next degree (in most men,) those that concern conjugall affection'.[28] Similarly, Nietzsche's noble class of rulers will still have 'duties towards their peers'.[29] So there is wider agreement on the nature of man as a creature who is generally concerned in some way about the welfare of at least some other persons. This means the challenge is not the same as that attempted by Bernard Williams, for example, in trying to bring into the moral realm someone for whom morality has no pull at all.[30] The immoralist has an interest in morality, but understands moral concepts differently, applying good and bad inversely to the kinds of actions and states to which we might conventionally apply them. So we are dealing with someone who is ostensibly already within the moral realm.

We can now return to the claim that we begin the journey into morality on grasping the punishment or reward we receive for actions which do not relate to our own self-interest. We see a look of anger or disgust which shapes our feelings about certain actions. In many cases these responses are directed at us personally, but we also witness the way others are treated. The child who sees her mother acting compassionately towards her sibling will be influenced on seeing this, just as she will on receiving a compassionate response to her own needs. The internalised other is constructed initially out of the simplest components - cries of help, looks of anger, displays of care - but these simple components are general features of the way people are, out of which a new set of responses is created in opposition to self-interest.

Still, why should we give greater priority to compassionate behaviour in understanding morality? Why not insist on rewarding the child who lashes out instead of punishing her, showing indifference to pain instead of responding in a caring way? Why should these arguments count as anything more than a sentimental appeal for conventional morality? It is here that the opening emerges for a more compelling response to the immoralist. If our morality reflects the way we are, then our moral theory must take account of characteristics such as our capacity for love and compassion. The challenge is to justify the claim that it is these traits rather than indifference to them which ought to

---

[27] See FN 28 above, for reference.
[28] Thomas Hobbes, *Leviathan*, 204.
[29] Friedrich Nietzsche, *Beyond Good and Evil*, § 260.

shape the content of the internalised other, where the 'ought' is not one rooted in the morality we are trying to justify, but in a naturalistic psychology which the immoralist has also accepted as the standard by which a moral theory can be judged. Nietzsche sees psychology as a means of debunking traditional morality, for 'Never yet has a *deeper* world of insight been opened to bold travellers and adventurers; and the psychologist who makes this kind of 'sacrifice' ....may demand at least that psychology be recognised as the queen of the sciences'.[31]

It is worth pointing out here that any of the major, conventional positions in moral philosophy will have their own means of confronting the immoralist, with the concepts of duty or virtue as two obvious candidates for how one might fight this battle. As I think will already be clear, the position I shall be defending is a form of sentimentalism akin to that of Hume or Schopenhauer, and either one takes emotions such as compassion and sympathy for the suffering of others as good, or else as the symptom of sickly modern European man. As I said earlier, the immoralist differs from the egotist in starting out with a concern for others, but has also advanced the claim that the kind of beings we are ought to determine our morality. The conclusion she derives is that the satisfaction of personal desires through their unfettered expression must be allowed to determine the content of our morality, but this is too quick. The kind of beings we are surely demands precisely the opposition conclusion. We prefer less pain rather than more. We are happier in love rather than in hate, in friendship rather than conflict. If someone strikes us, it would just be odd to reward the person who does it, or to punish someone for showing compassion for those we love, and a morality founded on such responses just seems bizarrely at odds with what we are at a profound level.

At this stage, if we are to turn to psychology (as Nietzsche demands), then utility is surely the better indication of fundamental drives than will-to-power or a striving for creative self-fulfilment. To praise pity is not to endorse a disguised form of contempt. As Philippa Foot says, 'Thinking of ordinary, unpretentious men and women who seem to find special happiness in working for the relief of suffering, one must surely find Nietzsche's dismissive views on compassion rather silly.'[32] Given our capacity both for love and kindness on the one hand, and for a need for others during our life, then a morality which reflects these qualities just looks like a morality which corresponds more closely to our basic needs than immoralism. The suggestion here is that utility underpins the content our morality (although I shall suggest later that this does not lead to a reduction), and is well-placed to do this given our preferences for happiness over misery. To turn to Foot once again,

---

[30] Bernard Williams, 'Egoism and Altruism', reprinted in *Problems of the Self*. Williams' argument addresses principally the issue of whether or not we can use reason to convince someone for whom standard moral considerations have no force to act in a moral way. I shall deal with this issue the next chapter, on moral reasoning.

[31] Friedrich Nietzsche, *Beyond Good and Evil*, § 23.

[32] Philippa Foot, *Natural Goodness*, 107.

'[l]ove and other forms of kindness are needed by every one of us when misfortune strikes, and may be a sign of strength rather than weakness in those who are sorry for us. We may reasonably think, moreover, that charity makes for happiness in the one who has it, as hardness does not.'[33]

Perhaps there is a further card the immoralist can play. I have suggested we are happier in love than in hate, in friendship rather than conflict. But are there not more salubrious sources of utility? Gore Vidal once remarked 'every time one of my contemporaries succeeds, a piece of me dies.' Is there not a sense in which sentiments such as this or *schadenfreude* reveal a darker side of human conduct which may also constitute elements of a happier life? If utility is to underpin morality and such responses are sources of utility, then why should they not count as virtuous thoughts?

The answer to this resides in the conceptual structure of the internalised other, which is not simply a metaphor for a recognition of the presence of others, or for the usefulness of others in providing a happier life for me. It can best be understood as standing in for the *claims* of others which come to partly constitute our own identity. In the scenario used earlier of the mother who darts back the look of anger at the child who strikes her, it is her claim not to be harmed which will come to shape the internalised other of the child, and these are the claims of a loving, caring, suffering being - someone typically human. Sentiments such as jealousy or *schadenfreude* just couldn't be ones which match up to the needs of those who are the objects of such emotions, and a full account would locate them in the realm of personal desires, perhaps ones which are shaped by the wider community, but not distinctive of the demands of those around us which determine the content of what could count as a moral response in the way I have characterised such responses here. The key point will be that once we characterise morality as a phenomenon which is rooted in the claims of others, then what others are will place a conceptual limit on what could count as being a moral response. Compassion is a response to the need of another person. *Schadenfreude* has another person as its object, but does not constitute a response to their claims or needs, and reflects primarily the content of our self-interested view on the world, where others may play an instrumental role in satisfying our desires of the world being the way we want it to be.

The immoralist starts with a concern for others, but then demands that this concern be restricted in such a way that aspects of who and what are we must be suppressed. Hume's naturalism is not only more attractive from the standpoint of conventional morality, but also theoretically more satisfactory than Nietzsche's because it relies on a more rounded account of what we are, and is therefore more defensible according to the standards of a naturalistic psychology. I also hope to have undercut the line that my appeal to utility is itself a product of the morality I am defending. This is the sort of line Nietzsche presented when he argued, 'The force of moral prejudices has

---

[33] *ibid.* 108

reached far into the most spiritual world, a world apparently cold and without premise – and it has obviously had a harmful, inhibiting, blinding distorting effect.'[34] Hobbes held a similar view of the ultimately bleak nature of human character, 'The Passions that encline men to Peace, are Feare of Death; Desire of such things as are necessary to commodius living; and a hope by their industry to obtain them.'[35] But if the empirical evidence on compassion is right,[36] then these claims seem vulnerable, and it is conventional morality which can draw on psychology to justify itself as the most appropriate way for humans to live, given the way we are. Compassion is a natural potential in all of us, and not an invention of the weak to surreptitiously overcome those who are naturally stronger than they are.

This line of argument also provides an account of how the defender of conventional morality will characterise the immoralist. The conventional moralist will perceive the world in a way which is conditioned by her capacity for compassion and benevolence. Where the immoralist differs is not solely on a theoretical point, but also on a perceptual one. To see pain as good is to see the moral world differently. There is an awareness of the moral relevance, but a lack of feelings with regard to pain which condition the perceptions of the agent whose morality is in line the psychological requirements that the immoralist herself accepts as the criteria for judging the validity of the position. This position can be justified in terms of our psychology, but not if we see compassion as good *per se*, for then we are vulnerable to the counter claim that we are simply in thrall to traditional morality, and if our attitudes are all we have to fall back on then there is no means by which to show our attitudes are more worthy of being deemed moral then those of the immoralist. Even the deep-rooted preferences I referred to earlier can be cast as the bleating of a modern European herd animal. But the internalised other places a different challenge before the immoralist. I have suggested that emotions such as compassion, sympathy and pity are not exclusively the products of morality, but emotions which are naturally occurring in us. We start out with these elements, as against acquiring them as a mutation of our hatred of the strong.

If these arguments go through, then the conclusion that sympathy and compassion, rather than self-assertion provide a more appropriate content for the internalised other follows from this, for it is these capacities which best meet the various challenges I have set out. They give content to a concern for others which is characteristic of the moral, and they do justice to the intuition that there are certain moral positions which simply have to be ruled out in the name of a proper morality. They also reflect at a deep level the kind of creatures we are, providing explanatory force as to precisely why such qualities should come to dominate our moral lives. Once we accept the claims

---

[34] Friedrich Nietzsche, *Beyond Good and Evil*, § 23.
[35] Thomas Hobbes, *Leviathan*, 73.
[36] This evidence will be set out in detail in the next chapter.

that sympathy and compassion are part of what we are, and that such qualities contribute to relationships such as friendship or love in which we are more likely to find happy and fulfilling than hatred and enmity, then a morality based on such qualities is one which we can defend as something which emerges out of our humanity.

At this point, it is worth emphasising that I have focused the discussion purely at the most basic level of which sorts of responses shape our feelings when we begin to enter the moral realm. None of this should be taken to suggest a reduction of morality to the simple categories of pain and pleasure. My aim is simply to sketch in the broadest outline the conceptual parameters of the internalised other as it first forms in us. I am claiming that a rounded view of what we are will inevitably undercut the immoralist's views, whom I have described as lost in the moral world by virtue of a failure to acquire the feelings which guide the properly moral person. The violent action towards those around her have never been sufficiently checked by the look of anger, or she has never internalised the model of care and compassion which leads one to concern about the pain or suffering of another. She has, in other words, only partially crossed the initial threshold of the moral realm, and the admiration of a world in which suffering is good or indifferent is not an alternative moral view, but the view of someone standing on the fringes of morality whose perceptions are restricted in virtue of her distorted emotional development.

In light of these claims, consider Nietzsche's yearning for an 'extra-moral' stage in human history where the will-to-power will no longer be inhibited by the pity with which Christianity has emasculated it, and new leaders 'under whose new hammer and pressure the conscience would be transformed into steel, the heart into bronze'.[37] He demands a suppression of the very emotional responses which I have characterised as fundamental to a proper morality. The responses from important others, responses which are characteristic of any human being would be repressed in order to toughen the conscience into steel, and the heart into bronze. On the account I have offered, he is demanding a suppression of certain aspects of what we are so that the strongest can rise to dominate. As such, he demands a shaping of the structures which constitute the internalised other in the ways which ignore fundamental tendencies both to need others and to respond caringly to their need of us. In the final analysis, Nietzsche's naturalism opposes the very elements of our natural make-up which Hume and Schopenhauer condone, the ones which make our morality reflective of our humanity.

I have characterised morality as the positive construction of an internalised other built out of certain natural capacities which come to restrict others. The impulse of a wider concern for others is set against the satisfaction of self-interested desires. The idea of the internalised other enables us better to defend much of the intuitive appeal in a sentimentalist account of ethics with regard to compassion as the basis of our moral lives. As they stand, I have suggested that the accounts offered by Hume and Schopenhauer are open to the criticism that there is a conceptual step missing in the move from benevolence or compassion to our moral concepts. This means that an opponent can either point out there is no logical contradiction in the claim that these sentiments are bad, or else claim that we should just see these as a set of attitudes which many see as good, but ought not to. In either case, we need a further argument to defeat the immoralist. My claim is that by focusing on what kind of beings we are, and a wide range of natural impulses such as caring for others and the wish to avoid pain, we can see that an internalised other who embodies these impulses in the form of value is closer to what we are.

These arguments indicate that a proper morality is one which is consistent with the kind of beings we are, and this means a morality which involves a concern for the claims of others which form part of our perspective on the world. When we talk of the claims of others, this means what they need and what we can do in response to those needs. It is compassion which constitutes the capacity in us to respond in these circumstances. A legitimate response to the need of another person must take account of what that person is, and her status as caring, suffering, loving being. Any morality which fails to conceive of persons in this light cannot be one which provides responses to needs, and therefore cannot be a proper morality. There are further obvious problems which arise at this stage, which will be the subject of the next three chapters, but which need to be signalled now. Firstly, we need to be able to say what counts as a need which we can press legitimately as a claim on others. Secondly, how far must we go in responding to the needs of others in order to satisfy the demands of morality? Both of these questions require response which require a fuller account of the nature of compassion, an indication of how reason be used to understand the limits of our legitimate needs and obligations, and how justice can regulate the extent of each in light of our conclusions to these questions.

There is, however, one further point of key importance which emerges at this stage. If this theory is right, then morality must be seen as irreducible. The internalised other is an aspect of our identity which is constructed out of simpler elements, but one cannot arrive at an understanding of the moral self by breaking down this aspect of our identity into its constitutive parts. If one is within

---

[5] Friedrich Nietzsche, *Beyond Good and Evil*, § 203.

the moral realm, then one's view of the world is already shaped by the claims of others which have become part of what we are. Insofar as one could imagine someone whose identity contains no trace of the claims of others, they simply could not have any understanding of what it is to arrive at a moral judgement. There could be no more than an awareness of certain behaviours and dispositions in others, but no understanding of what is like to hold one of the judgements from which this behaviour stems - they would not have acquired the concepts which are made available only once we have the perspective of the internalised other. They would know how people behave when they are being what is called 'moral', but they would not know what it is like to see the world in a moral light. To characterise the moral perspective as one which is dependent upon the claims of others, is to characterise it as a perspective which is other regarding, and we thereby obstruct any move to reduce it ultimately to questions of self-interest. Utility, as I have suggested, shapes what the internalised other may become, but it is not identical with morality. It is our preference for utility and our capacity for our feelings to be shaped by others which combine to make the internalised other and moral identity possible. In other words, the moral self is something distinctive which is born out of what we are. Curiously, we find some support for such a view even in the thinker to whom I have been most opposed in this chapter. Even for Nietzsche, it is because of the constraints which morality places on us that 'something emerged, and emerges that makes life on earth worth living'.[38]

---

[38] *Ibid*, s. 188.

6. Compassion

Introduction

The conclusion of the last chapter was that the view that compassion is the basis of a proper morality is one which can be defended in light of certain key considerations about the nature of human beings. In particular, the need we have of others which is revealed through suffering and our capacity to respond to such needs when others express them, suggest that compassion is a basis for a morality which makes sense in light of the kind of beings we are. I have also suggested that the moral perspective is one which is irreducible, in that it involves the acquisition of a distinct range of concepts which become available only by means of the development of what I have called the internalised other. To have acquired a moral perspective in this way is to take on a particular way of seeing the world, and not to have such a perspective would rule out the possibility of understanding events around us as a moral agent. The moral perspective involves the use of certain feeling-laden beliefs and perceptions which are indispensable in coming to what we would we would call a moral understanding. Only the emotional responses which are part of this perspective make possible the meaningful use of moral concepts.

This much has already been argued for, but if a proper human morality is to be one grounded in compassion - one where compassion characterises the moral perspective - then we need to come to a more detailed understanding of the nature of compassion in order to give a clearer picture of what such a moral perspective must involve. The aim of this chapter will be to address in greater detail what we mean by compassion, how we come to be compassionate, and what specific implications this has for the current theory.

What is Compassion?

There are at least three reasons why it is difficult to arrive at a precise understanding of this term. The first is that there are so may other concepts which appear to be so closely related – sympathy, empathy, benevolence and pity in particular. Secondly, many of these terms are used as if they were interchangeable. In a paper on *empathy*, one example of an early manifestation is described as follows, 'At nine months, Hope had already

demonstrated strong *sympathetic* responses to other children's distress' (italics added).[1] Finally, there seem to be important distinctions in the scope of compassion, as it is discussed by various philosophers. For Butler, it is 'the sight of a man in misery' which raises our compassion,[2] suggesting that it is suffering alone which is the object of our compassion, but for Schopenhauer the emphasis on suffering is extended to include the 'weal and woe' of another, indicating that not simply the removal of suffering, but also a wider sense of benevolence has a part to play.[3] In the first instance, it will therefore be useful to set out what I shall mean when I refer to compassion, and how I shall understand the distinctions with regard to other closely related concepts.

I shall take compassion to be characterised by being moved by the suffering of another. Being moved should be understood as involving the feelings associated with compassion, which are combined with the belief or perception that someone is suffering, and the desire to alleviate the suffering. That suffering should be the object of a compassionate response seems to follow from any conventional understanding of what compassion is, which seems to be a response to 'A person in a negative condition, suffering some harm, difficulty, danger'.[4]

One crucial point to be noted from the outset is that the concept of 'suffering' is to be understood in a very broad sense which stretches beyond physical pain. I will take suffering to include psychological distress and also deprivation or a sense of loss incurred by the victim. In order to highlight the reasons for this, consider the possible impact of one of the more notorious dimensions of the Taliban's rule in Afghanistan. Imagine a woman who, as an adolescent had dreamed of becoming a philosophy lecturer, but who had been denied access to education and found herself pressurised into the more conventional life of a traditional wife and mother. Let us say she expresses resentment and regret at having been denied an education, but seems to lead a moderately happy life. We may well feel compassion on the grounds that she has been denied an opportunity to flourish in the way she wanted and has missed out on certain possibilities which we may feel she ought to have had, but this isn't a response to physical suffering, and we need a

---

[1] Ross A Thompson, 'Empathy and Emotion Understanding: the Early Development of Empathy', in N Eisenberg and J Strayer edd, '*Empathy and its Development*'.
[2] Joseph Butler, 'Dissertation on the Nature of Virtue', in D D Raphael ed, British Moralists 1650-1800, vol 1, 380.
[3] Arthur Schopenhauer, *On the Basis of Morality*, 144.
[4] Lawrence Blum, 'Compassion', in *Moral Perception and Particularity*, 173.

broader understanding of what we can perceive as suffering in order to explain our reaction. The range of circumstances which can occasion compassion seems to go beyond paradigm cases such as physical pain or an acute sense of psychological pain on the part of the person with whom we are concerned, and for this reason a wider concept of suffering is intended to broaden the range of objects to the point where they coincide with the range of our compassionate responses.

It is also worth elaborating further on what it is to be moved by the suffering of another, particularly in order to distinguish it from two mistaken conceptions of what compassion is. Firstly, when we say we are pained at another person's suffering, this does not mean that we experience the same distress as them. Humean sympathy seems to work like this.

> 'Tis indeed evident that when we sympathize with the passions and sentiments of others, these movements appear at first in *our* mind as mere ideas, and are conceiv'd to belong to another person, as we conceive any other matter of fact. 'Tis also evident, that the ideas of the affections of others are converted into the very impressions they represent, and the passions arise in conformity to the images we form of them."[5]

If we take this as a view of compassion (which Hume did not), then it looks as if I perceive someone as grieving at the loss of a loved one, and I then experience grief as well. But this is surely no more accurate than saying that if I feel compassion at the sight of a car crash victim I experience the same sort of pain as they do. This phenomenon is better described as an emotion 'resonating', whereby I might see you are sad at the tragic conclusion to a film and I become sad myself in response to the same stimulus.

A closely related error to thinking that compassion involves experiencing the same emotion is assuming that compassion involves responding to the same object. This also gets it wrong, in that although the suffering of two people may have the same object, this doesn't indicate compassion on the part of one for the other. If I am sitting with a compatriot watching the English batsmen wilt before the Australian bowling attack, we will each suffer, but this isn't yet grounds for saying that either of us is responding compassionately towards the other. Compassion takes as its object the suffering itself,

---

[5] David Hume, *A Treatise of Human Nature*, 319.

not the cause of it. If I feel compassion for a friend grieving at the loss of his mother, it is the suffering he endures in the form of his grief which is the object of my compassion, not the death of the mother. With these initial comments in place, it is worthwhile distinguishing between the concept of compassion and other similar concepts.

Related Concepts

In order to provide maximum clarity, I shall start with those concepts which are most distant and work in towards those which appear closest to compassion as I have described it.[6] Perhaps the most obvious distinction comes with empathy. As I said earlier, this often appears in a way which looks synonymous with compassion, but is best understood as a quite separate phenomenon. The *Chambers Dictionary* defines empathy as 'the power of entering into another's personality and imaginatively experiencing his experiences: power of entering into the feeling or spirit of something.' The reason for treating this as importantly distinct from compassion is that one can empathise with the pain of another without being moved by it in the appropriate way or even in any way. Martha Nussbaum points out that a torturer may empathise in order to develop methods which will have maximum effect on her victim, meaning that empathy might not merely fail to deliver up a compassionate response at the prospect of someone suffering, but may even prompt satisfaction at uncovering a more effective means of causing it.

Nussbaum also points out that empathy can nevertheless be a route to compassion, in that we can imaginatively construct what we believe someone will feel in a particular set of circumstances, and be moved by the prospect of their suffering in the imagined scenario. One possibility then, is that empathy may be a necessary component of compassion, which results in a compassionate response in certain people and not in others. This is a view which occurs fairly frequently among child psychologists, where compassionate response are often seen as tied to empathy. For example, Paul L Harris writes that,

---

[6] I shall follow closely Martha Nussbaum's discussion of the distinctions set out in Chapter Six of *Upheavals of Thought*.
[7] *Ibid*. 329

> 'For the moment, the important point to underline is that cognitive factors, particularly the ability to look at a situation from another's perspective, appear to be intimately connected with the willingness to offer comfort.'[8]

Such a view seems rather unlikely given the age of the infants who first respond in what one might term a proto-compassionate manner. It is difficult to believe that an infant of only eight months is undertaking such a complex calculation, and this looks more like an over-complication of the actual process. It seems much more likely that such early cases involve the infant perceiving the suffering of another, and being moved on grasping that the person is suffering, as against considering how it would be for the infant herself to undergo such an experience, understanding how unpleasant it would be, and then responding. Crying, screaming and certain facial expressions are perceived by the infant, and the suffering provokes the distress. The richness of our compassion needn't be present in our earliest experiences of it, but can develop later.

A further argument which casts doubt on the view that empathy plays any necessary role in compassion is that we might plausibly feel compassion for those with whom we cannot empathise. If I am moved by the sight of my beloved dog gradually slipping away, I might sincerely assert feelings of compassion for him, but this can hardly because I have the power to enter into his personality and imaginatively experience what it is like for a dog to die. It is the perception of suffering which is doing the work in this case, regardless of my inability to empathise.

If empathy is neither necessary nor sufficient for compassion, then pity seems somewhat closer to an understanding of compassion. Nussbaum's view is that compassion involves not simply the understanding that someone is suffering, but a feeling at some level that this is a bad thing. Pity seems to come closer to meeting this requirement than empathy does, in that it involves 'a form of conviction that someone else is in pain',[9] where the form includes the impression of this pain being bad or unfortunate. The possibility of pitying someone and being indifferent to their suffering seems much more problematic than is the case with empathy. That said, pity also differs from compassion in certain key respects. Firstly, it has come to have connotations of condescension. The thought of

---

[8] Paul L Harris, *Children and Emotion*, 41.
[9] Ludwig Wittgenstein, *Philosophical Investigations*, § 287

being pitied is unpleasant to us because it brings with it a certain humiliation – being pitiable can stretch beyond being in need of help or deserving of help, into being lower or even repellent.

Probably the best example of the complex mix which goes into pity comes in Nietzsche. On one hand, he extols pity on the part of the strong, yet also finds in it a more poisonous side. We are told at one point that 'a man who is by nature a *master*: when such a man feels pity – well! *This* pity has value. But why should we bother about the pity of those who suffer! Or of those who may even preach pity![10] Even if we reject Nietzsche's view of pity as virtuous in the strong and irrelevant or unimportant in the weak, he is surely right about the potentially double-edged nature of pity as a response which can bring with it a form of contempt. Perhaps the best explanation for this is that pity seems to have wider scope than compassion. Compassion is directed towards someone's suffering whereas pity can be directed much more widely. One can pity someone for who or what they are, as in cases where a religious bigot might pity someone for holding a conflicting set of religious views, or for having chosen a particular profession or style of life. In stretching beyond a concern for our suffering, which we ourselves will usually wish to end, and into a value judgement of a wider range of qualities which we may hold much dearer, pity can include a disdain or lack of respect for aspects of who we are which we value.

It is this aspect of pity which allows Nietzsche to present 'the priestly class' as the 'greatest haters in history'.[11] The preoccupation with not simply the suffering of others, but also with a wider moral scheme can be smuggled into pity. The more tightly focused nature of compassion makes it less amenable to being co-opted into this sort of project. This doesn't amount to the claim that pity *must* entail contempt or opposition at some level – pity can be as specifically concerned with someone's suffering as compassion – but the fact that it presents this possibility opens up a conceptual gap with compassion.

There is one further point at which this gap appears. We talk about 'self-pity', in which we feel sorrow concerning our own situation. But if the account of compassion I am giving is correct, compassion for oneself is impossible by virtue of the fact that it is

---

[10] Friedrich Nietzsche, *Beyond Good and Evil*, § 293.
[11] Friedrich Nietzsche, *Genealogy of Morals*, § 7.

necessarily directed at the suffering of another. I suggest that it is simply part of the meaning of compassion that it is directed towards another, thereby further separating it from pity, on which there is no such constraint.

We come a step closer to compassion when we consider benevolence, indeed one might even ask whether compassion could not be subsumed under benevolence. If we return for a moment to the example of the woman living in Afghanistan under the Taliban, could one not explain our attitude towards her as an example of the sentiment of benevolence described by Hume as a 'tendency to promote the interests of our species, and bestow happiness on human society'.[12] One could argue that it follows trivially that an attitude of benevolence entails the wish to see someone's pain ended. At this stage, we have the additional judgement which is absent from empathy, and as yet no reason to believe there is the contempt which can be concealed within pity.

Even if being the recipient of someone's benevolence may make us less uncomfortable than being the recipient of someone's pity, there are still good reasons for emphasising the distinction with compassion. Like pity, benevolence seems to go further, in that being moved by someone's suffering is merely one form benevolence, and certainly doesn't exhaust the possibilities. The presence of suffering is clearly unnecessary for benevolence, in that a millionaire who leaves his fortune to his wealthy son may well act out a sentiment of benevolence, and certainly benefits him, but needn't do so in a way which can be understood as alleviating any form of suffering. In addition, having said that benevolence appears to lack the potentially unpalatable side of pity, there is nevertheless a side to it which has provoked concern amongst many liberal thinkers. In our aim of benefiting someone, we need not be restricted by the more modest task of removing a source of suffering, but might engage in what we sincerely believe to be a project of helping them to become someone better. Benevolence naturally stretches beyond a concern for freedom from suffering into a concern for the freedom to be a particular sort of person or lead a particular sort of life.[13] It aims at promoting the good, one means of which may be removing the bad. But compassion is concerned principally

---

[12] David Hume, *An Enquiry Concerning the Principals of Morals*, 20.

[13] As my aim here is purely to tease out the conceptual distinction between benevolence and compassion, I shall not enter into the debate between proponents of negative and positive liberty. The classic statement of the distinction between the two comes in Isaiah Berlin's 'Two Concepts of Liberty', with alternative views presented by Gerald MacCallum in 'Negative and Positive Freedom'

with what the victim suffers from, as against what we think they would benefit from. This means that the voice of the victim is of crucial importance, which needn't be the case with benevolence. With compassion, we respond to the actual suffering, or to the prospect of some future suffering where the victim herself suffers, and this is the ultimate focus of our response. With benevolence, the focus of our response is benefit, and the conceptual distinction means that compassion and benevolence can even be opposed to one another. I can reach the view that suffering strengthens you in ways which are beneficial, and override your protests in exercising my generosity in pursuit of this goal. As such, benevolence might endorse suffering and oppose a compassionate response to it.

Sympathy is perhaps the term which comes closest to capturing the attitudes which I have suggested lie at the heart of compassion. Both are normally activated on apprehending that someone is in difficulty, and involve being moved by that perception. Their proximity is such that they are surely interchangeable in many contexts, so that when Hume talks of no qualities being more entitled to 'the general approbation and good-will of mankind' than those which proceed from a 'tender sympathy with others',[14] one could easily imagine substituting 'sympathy with' with 'compassion for'.

Nevertheless, there are perhaps two points at which important distinctions can be drawn. The first is that sympathy can shade into something like empathy in a way in which compassion cannot. One can be sympathetic to the views of a classical theist who sees love of mankind as a her guiding principal in her dealings with other people. Sympathy here indicates an affinity or agreement at some level in the way one feels about a particular subject. No suffering need be involved. Conversely, compassion can indicate an intensity of feeling which sympathy may lack. Consider how best to capture the attitude of Achilles when he responds as follows to the pleas of Priam for the return of Hector's body. "Poor man, how much you've borne – pain to break the spirit!"[15] There is a sense in which describing Achilles' reaction as one of sympathy seems to understate the extent to which he is moved by the plight of Priam, and compassion seems better

---

and Charles Taylor in 'What's Wrong with Negative Liberty'. All are reprinted in David Miller ed, *Liberty*.

[14] David Hume, *An Enquiry Concerning the Principals of Morals*, 18.

[15] Homer, *The Iliad*, Bk 24, line 605.

suited to convey the force our emotional responses in the face of such powerful circumstances.

I have suggested that there are distinctions between these concepts which emerge under certain circumstances, but this is not to imply that these difference are hard and fast, or that they are always obvious. Achilles could be said to feel pity for Priam, and compassion can be experienced as one of the 'calm passions' which guides our behaviour without necessarily involving the kind of intensity present in the scene from the *Iliad* described above. But these kinds of fine-grained distinctions are important if we are to describe accurately the particular aspect of our moral lives which is being picked out. If one is to express the phenomenology of compassion, then it seems to be captured in the thought 'Let the suffering stop!'. Each of the alternative concepts discussed above varies somewhat from this basic impulse of compassion, and it now remains to explore to a greater extent what is involved in this specific emotion.

The Content of Compassion

The previous chapters have set out in detail the characteristics of an emotion, and it is now a question of explicating compassion within this context. If we retain the account of emotion developed earlier as a feeling and a belief or perception combined in consciousness, then compassion can be seen as the belief or perception that someone is suffering combined with a feeling towards the suffering and a desire to alleviate it. More now needs to be said about the nature of the beliefs which are involved.

Nussbaum follows Aristotle in arguing that the cognitive side of compassion involves three separate elements.[16] Firstly, there must be the impression that the suffering is serious rather than trivial. Secondly, the person must not be seen as having deserved the suffering. Thirdly, the person who is experiencing compassion must be aware that they are seeing someone suffer in a way which may reveal their own vulnerabilities. I shall argue that only the first of these is a necessary condition for the cognitive side of compassion.

---

[16] See Martha Nussbaum, *Upheavals of Thought* 306, and Aristotle, *Rhetoric*, in *The Complete Works of Aristotle*, ed Jonathan Barnes.

One might argue that the idea that the suffering must be viewed as important represents an over-intellectualisation of what goes on. Surely we simply grasp that someone is distressed and respond compassionately. This may well be true in the case of infants who cannot distinguish between important and unimportant forms of distress, but there is every reason to think that as we develop we become much more discerning. Consider for example whether or not we would feel compassion on witnessing the distress of the American undergraduate on a study-abroad programme in Europe, portrayed in Jonathan Franzen's *The Corrections*.

> Cheryl said to Tiffany: "So my dad's like, you've got to sublet if you're going to Europe, and I'm like, I promised Anna she could stay there weekends when there's home games so she can sleep with Jason, right? I can't take the promise *back* – right? But my dad's getting like all bottom-line, and I'm like, hello, it's *my* condominium, right? You bought it for *me*, right? I didn't know I was going to have some stranger you know, who, like, *fries* things on the stove, and sleeps in my bed?"
>
> Tiffany said: "That's so-gross."
>
> Cheryl said: "And uses my pillows?"[17]

It seems entirely plausible that Cheryl is genuinely distressed at the thought of renting out her home, yet we may well find ourselves less than concerned. The question of how we come to see some suffering as worthy of compassion and some not is one which I shall return to in a later section of this chapter, but for the time being it is simply worth emphasising that compassion involves a level of discrimination by the agent as to the importance of the suffering, with trivial forms of suffering unable to provoke the emotion.

What then of Nussbaum's claim that we must see the suffering as undeserved. She argues that 'Insofar as we believe that a person has come to grief through his or her own fault, we will blame and reproach rather than having compassion. Insofar as we do feel compassion, it is either because we believe the person to be without blame for her plight

---

[17] Jonathan Franzen, *The Corrections*, 456.

or because, though there is an element of blame, we believe that her suffering is out of proportion to the fault."[18]

It seems fair to say that our compassion may be diminished or ultimately evaporate as a result of the agent being responsible for bringing the suffering on herself, but it seems questionable whether the conceptual link between fault, suffering and compassion is quite as close as Nussbaum suggests. There seem to be too many cases where agents engage in activities when they are fully aware of the dangers, and yet we are moved by their suffering when things go wrong, despite their bearing a degree of responsibility for their plight. Those who suffer from drug addictions are surely capable of provoking our compassion even where it is tinged with reproach for the irresponsibility and selfishness of the decision to start down this path in the first place. And this isn't because we see the suffering as disproportionate, we surely just see it as bad that anyone should suffer like that, regardless of who is responsible. Granted, our compassion may be greater where the victim is entirely innocent, but there are certain points when we want to say that no-one should suffer like this, regardless of what they have done, and fault now becomes almost irrelevant in our consideration of whether or not they are worthy of compassion. To take an extreme example, one might feel compassion for the murderer on his way to the gallows.

A further problem for the view that there is a hard and fast link between fault and compassion is the way in which events after the fact can influence our response. If we take the example of someone who commits a terrible crime, and initially remains defiant and unapologetic, then their incarceration is something we may well see as something good. But if, after some years, we see in them sincere remorse for what they did, and what we take to be a proper understanding of the horror of their crime, is there not a sense in which the person now becomes someone for whom we may feel a compassion which was previously absent? We may begin to question whether or not it is right for them to be imprisoned for life, and to feel that they now deserve to be given another chance. If this is the case, then our responses can change independently of the degree of fault or suffering, which can remain unchanged. The remorseful person may suffer additionally as a result of seeing what they have done in a different light, but we don't start to feel compassion because their suffering has been increased. We start to feel

---

[18] Martha Nussbaum, *Upheavals of Thought*, 311.

compassion because we are moved by the plight of someone who is remorseful, but less so by that of someone who is unrepentant about having committed an action which we see as deplorable. If our view of someone turns to compassion where there is no change in either the degree of responsibility or the level of suffering, then the link between fault, suffering and compassion cannot be as tight as Nussbaum suggests.

One might even go further and suggest that we can feel compassion even where the punishment is deserved and the victim unrepentant. Consider for example the case of Pierre's duel with Dolohov in *War and Peace*. Dolohov is brash, sneering, and assumed by Pierre (wrongly as it turns out) to be having an affair with his wife. At a dinner, Pierre becomes increasingly antagonised by his rival's perceived condescension. 'I know him. He would find it particularly alluring to besmirch my name and hold me up to ridicule after I had exerted myself on his behalf and befriended and helped him.'[19] Later on, Pierre feels himself provoked beyond endurance when Dolohov proposes a toast to him. 'Here's to the health of all beautiful women, Peterkin – and their lovers!'.[20] A further slight results in Pierre's challenging him, and they fight a duel. Killing or wounding an opponent under such circumstances would be entirely consistent with, or even demanded by, the code of honour by which a Russian gentleman lived, and yet when Pierre surprisingly succeeds in wounding his opponent, he is mortified at the harm he has done. 'Pierre, hardly able to restrain his sobs, started to run to Dolohov'.[21] The knowledge that his actions might be formally justified counts for nothing at the sight of the suffering, and his response is a 'smile of compassion and regret'.[22] Even 'Dolohov the brawler, Dolohov the bully', so deserving of his fate according to the standards which commonly applied, becomes the object of compassion from the person who ought above all to see him as receiving the punishment he deserves.

What, then, of the claim that if I am to feel compassion for someone, I must be aware that I myself am vulnerable in the same way? Nussbaum takes the presence of 'similar possibilities' to be necessary, justifying this claim on the grounds that 'one makes sense of the suffering by recognizing that one might oneself encounter such a reversal.'[23] She argues that we come to understand the meaning of the suffering partly through

---

[19] Leo Tolstoy, *War and Peace*, 364.
[20] *Ibid.* 365.
[21] *Ibid.* 369.
[22] *Ibid.* 369.

understanding how it would affect us if we were to undergo it, thereby emphasising the close relationship between compassion and fear. This also helps to explain why it is that we don't feel compassion for just anything, in that only sentient beings capable of suffering share with us the possibilities which make us able to see that we could suffer in that way too.

Once again, there seem to be crucial elements of this claim which advance our understanding of compassion, but in the midst of an overall argument which is too strong. Consider the question as to whether or not the God of classical theism felt compassion for the Jews when they were enslaved by the Pharaohs. We would surely want to see compassion for the suffering of mankind as a consequence of the love such God would have for us, but his omnipotent, immaterial and eternal nature surely make it unthinkable that He might be enslaved by humans. Nevertheless, there doesn't appear to be any *prima facie* contradiction involved in the idea of God's being compassionate. In addition, there are cases where we feel compassion for people who are suffering in ways in which we couldn't. Babies born with a congenital illnesses from which they will die within a few months suffer from having their lives truncated in such a way which is impossible for us. Even though it is a contingent aspect of my life that I avoided such a horror, it is now an impossibility that it should ever befall me. Similar cases would be when childless couples, beyond the point when it is possible for them to have children, feel compassion for someone who has lost a child, or when we feel compassion for a suffering animal.

Where Nussbaum's argument carries greater force can be seen if we loosen somewhat the requirement that each of us must be aware of our individual vulnerabilities to feel compassion. It is surely more convincing to talk of the way in which seeing others suffer, regardless of our own specific circumstances, alerts us to dangers and vulnerabilities of people in general. This may sometimes include aspects of our own condition, in which case compassion can make us aware of how fragile our own lives are. In such cases, compassion can play an epistemological role from which our knowledge of ourselves and others is increased, but this removes the requirement that we must share the potential vulnerability of those whom we see suffering.

---

[23] Martha Nussbaum, *Upheavals of Thought*, 316.

There is a further reason for being suspicious of the view that compassion relies on shared vulnerabilities between the compassionate agent and the person who is suffering. If the thought is that we feel fear on realising that the state in which we see someone else could be one in which I might end up myself, it calls into question the ethical credentials of compassion. If I help someone lying in the street because I personally fear such states, then this looks like an explanation of compassion which is reducible to self-interest. If my fear is a necessary condition in bringing me to see that I should help someone, in that I could not reach this conclusion without experiencing fear, then concern for the fate of others requires a healthy concern for myself if it is to get off the ground. This surely flies in the face of our understanding of compassion at a profound level, in that our compassion is directed specifically at the suffering of *another*, and doesn't rely on our own imagined suffering for its motivational force. Whilst I may be moved to help someone as a result of grasping how much I would dislike it if I were in their shoes, this is clearly distinct from compassion. If fear is doing the legwork here, there is a basis in self-interest. If not, we are back to empathy again as a route to compassion.

The point of these reflections is to emphasise how much rests on what is to count as important suffering. Factors such as blameworthiness, fault, degrees of suffering and a sense of our own vulnerability may have some explanatory role in some cases, but it is only when we understand how we come to discriminate between important and unimportant forms of suffering that we can understand how compassion works. The central structural relationship is seeing suffering in a particular light, and what is required is a better understanding of how it is that we come to 'see' in this way.

The Shaping of Compassion

If compassion is constituted by a belief or perception about the suffering of someone, a feeling towards that suffering and a desire to stop it, then the key to understanding when we come to see suffering in a compassionate light will be understanding how it is that our desires and our feelings towards suffering come to be directed in the way they are. The model I have proposed for emotions in general is derived from Davidson's concept of triangulation, and the corresponding process described within child psychology is

social referencing.[24] We now need to consider how this explanatory framework can be used to chart the development of our capacity for compassion. To be more precise, we need to understand how it is that we come to see some suffering as important, and other cases as trivial. I shall begin by offering an explanation which derives from the model set out in the previous chapter, and then offering evidence from empirical psychology to support this account.

What each of us starts with, is a capacity to be moved by the distress of another person. This capacity will develop to varying degrees according to our social environment. This means that the behaviour of the primary carer towards ourselves and towards others will shape the way in which we respond to the suffering of others. Learning to respond to different degrees of suffering evolves through seeing certain levels of suffering, such as physical pain, receive more intense concern than other forms of suffering. This sort of concern is expressed through activities such as physical comforting, and also through more subtle ways such as the tone of voice and the facial expressions of the carer. Lower levels of suffering result in much lower levels of concern from the carer, and the infant who observes this comes to see these sorts of suffering as less important or trivial. The input from which our feelings come to be directed in these ways is consistent and patterned. Similar levels of distress receive similar responses time and again, to the extent that we abstract from the examples.[25] The overall impact of this shaping of our feelings is that we develop into persons who distinguish between which levels of suffering are important. It is these patterns and the responses which they inculcate in us which are the norms that govern our behaviour in conjunction with reason.

If this model is correct, there are various ways in which it ought to be identifiable within the empirical evidence regarding our development. There ought to be a correlation between the behaviour of the primary carers in response to distress, and our emotional development. More specifically, in cases where the carer shows greater concern, this should result in turn in a greater sense of concern for the distress of others. Conversely, where a child receives less attention when they are in distress, or sees less attention being lavished on others when they are upset, this should induce greater indifference to other people's suffering.

---

[24] See 56-62 above.

[25] This is not to suggest that perfect consistency or reproduction of the same responses is required, any more than it is when we learn to speak.

Before considering the evidence it is worth offering a word of warning about problems within many of the empirical studies, which relates to the distinctions between compassion and other closely related concepts discussed at the beginning of this chapter. There is an overwhelming tendency to talk of empathy and sympathy as if they were interchangeable, to the extent that it sometimes becomes difficult to know which concept they are referring to. For example, Mark A Barnett tells us that 'A child who has been shielded from particular distressful experiences, or discouraged from displaying overt expressions of distress would presumably have a difficult time empathizing with a needy other whose predicament or emotional reaction is perceived as foreign or unfamiliar.'[26] In such cases, it is not clear whether lack of sympathy for the distress of others makes empathy more difficult, or lack of empathy makes sympathy more difficult, or both.

The reason for this tendency not to distinguish between the two may well partly stem from an apparently widespread assumption within psychology that we start out as egoists and acquire the capacity for sympathy starting from our selfish drives and travelling via empathy. For example, Eisenberg and Miller state that 'the child learns that behaviors that make other people happy or relieve their distress are pleasurable for the self as well. Thus prosocial behaviors are, in essence, self-reinforcing and frequently performed because of the child's experiencing others' affective states.'[27] I have already rejected the model of empathy being either necessary or sufficient for compassion,[28] and the same arguments will hold against empathy as a pre-requisite for sympathy in cases where it is sympathy for the suffering of others.

With these caveats in mind it is nevertheless possible to draw on the results from a wide range of research, and I shall argue that much of the evidence supports the explanatory framework for the development of compassion which I have set out. In the first place, what one might plausibly interpret as proto-compassionate responses come as early as

---

[26] Mark A Barnett, ''Empathy and Related Responses in Children', in N Eisenberg and J Strayer eds. '*Empathy and its Development*'.154

[27] Nancy Eisenberg and Paul A Miller, 'Empathy, Sympathy and Altruism: Empirical and Conceptual Links', in N Eisenberg and J Strayer edd, '*Empathy and its Development*'.

[28] See 110 to 111 above.

eight months of age.[29] This seems improbably early for an infant to have acquired the relatively complex skills associated with empathy, whereby they would have to come to a view as to how they would feel if they were in the place of a distressed person. The unliklihood of such a process being able to explain compassionate responses in children of such a young age is compounded by the observations surrounding the separate question of the development of a sense of self and other. If empathy requires an understanding at some level of what *another* is experiencing and how *I* would feel in their place, then it also requires a sense of self and other to be in place. But the evidence suggests that this sense does not usually emerge until a later stage of development than that at which infants are capable of reacting sympathetically to another's distress.[30] If this is the case, then even if empathy may be one route to compassion on certain occasions, it cannot be the only one.

How then does the empirical evidence support the positive claims I have made for the development of compassion according to the behaviour of those around us? Two initial observations which need to be addressed are that there is a wide variety of responses identified in children in response to others' distress, and that there is a good degree of consistency in the reactions of individual children.[31] The issue then becomes one of how we explain the reactions in each case, and whether or not they emerge as a result of environmental factors. There is evidence from at least two studies which support the model set out above. Carolyn Zahn-Waxler[32] *et al* argue that particular forms of behaviour on the part of mothers influence the future emotional responses of their children. For example, in cases where an infant has hurt someone, and the mother both comforts the victim and explains in a 'moral tone'[33] both that this is wrong, and why, the child will tend to go on to comfort others when they are in distress. Similarly, when an infant is not the agent of the other's distress, but merely a witness to it, and the mother offers what one might think of as a concerned response to the victim, the infant witnessing this is more likely to develop strategies for comforting those who are in distress.

---

[29] See Ross A Thompson, 'Empathy and Emotional Understanding: the Early Development of Empathy', in Eisenberg and Strayer eds, *Empathy and its Development*, 119-45.
[30] Paul L Harris suggests that we first develop a sense of self and other at around two years of age. See *Children and Emotion*, 56.
[31] *Ibid*, 32
[32] Carolyn Zahn-Waxler, M Radke-Yarrow & R A King, 'Child Rearing and Prosocial Dispositions Towards Victims of Distress', in *Child Development 50*.
[33] This is Harris' interpretation in *Children and Emotion*, 36.

A second and more disturbing source of evidence comes from a study on children from abusive backgrounds. Mary Main and Carol George compared groups of children from similar social backgrounds, some of whom had been victims of maltreatment within their families.[34] The findings were that abused children did not exhibit obvious concern when confronted with other children who were upset, and tended to react aggressively, sometimes even attacking the other child. The children from homes where there was no record of abuse tended either to make a simple attempt to comfort such as patting on the head, or often displayed a much deeper concern.

Similar patterns seem to emerge from both studies. Where the parents exhibit concerned, comforting behaviour in the face of distress, and try to offer reasoned explanation, the infant will tend to develop into someone who is more concerned with others' distress. Children who have been abused tend to react not simply with indifference, but even with outright hostility to the distress of others. It would be rash to suggest that this material offers conclusive evidence for the theory I have advanced, and many of the studies in this field stress not only the relative youth of the research into children's emotional development, but also the presence of various interpretations of the data. Nevertheless, there are three factors which emerge from these studies which provide support for the outline I am defending here. Firstly, it is difficult to sustain the view of our emotions, including compassion, being somehow 'hard-wired' into us from birth. Such a position would posit an ability to respond to suffering with which we enter the world, or which develops in much the same way our powers of sight or hearing do. It is just there at birth, and a standard range of external stimuli will allow it to develop. The empirical research will support the claim that this is a natural capacity for which external stimuli are necessary if it is to develop, but also indicates that both what we perceive as suffering and which suffering is important is determined to a greater degree by social environmental circumstances than, for example, colour perception.[35] A second point closely related to this first one is that there is strong evidence to suggest our social environment is of crucial importance in shaping our compassion. How compassionate we are, how powerful our responses are to others' suffering, and whose suffering we are moved by are all susceptible to social influence. And finally, the early stage at which

---

[34] Main and George, 'Responses to Abused and Disadvantaged Toddlers to Distress in Agemates: A Study in the Daycare Setting', in *Developmental Psychology 21*.
[35] See 58-59 above for further discussion of this.

compassionate responses have been identified indicates there is strong reason to consider it as a direct response to others' suffering, rather than one which is routed through empathy.

I have focused exclusively on compassionate responses in infants in order to support my claims about its direct responsiveness to the suffering of others, but there is a danger of underestimating the deeper nature of compassion which we experience as adults. Whilst compassion may start out life as a relatively crude response to the suffering of someone in the immediate area, what I have called proto-compassion, it comes to take on a much richer form.

Once we have language, we can be moved by descriptions of the plight of others who are distant from us. We acquire a conceptual resource which *deepens* our emotional lives in comparison to the cruder emotions we experience as pre-linguistic infants. If an infant experiences compassion, it is a not a response which she can reflect upon. If we as adults are moved by the plight of someone close to us grieving at the loss of a family member, language allows us to consider how awful the loss is, judge how best to offer comfort, contemplate the value and significance of our own friends and family. This in turn may provide the basis for further emotional responses. We might suddenly come to reaffirm our feelings for those around us in light of our reflection, or perhaps be brought closer to the person who has endured the loss on understanding the full significance of such a tragedy. One way of phrasing this is to say that although infants may experience compassion, they cannot recognise it as compassion and reflect upon the nature and significance of their experience. It is language which opens up these further aspects of our emotional lives, and it is in describing these sorts of phenomena that we come to see what is meant by having a greater depth to our emotional responses towards others. Raimond Gaita discusses an analogous distinction when comparing the lives of animals to those of humans. He writes that,

> Only human beings (of the beings we know) have an inner life. That is because only human beings can reflect on what happens to them and take an attitude to what happens to them because of such reflection. An animal can suffer, but it cannot curse the day it was born; an animal can be afraid, but cannot be ashamed of its fear and despise itself; an animal can be happy but it cannot be joyous; an

> animal cannot give of its substance to certain pursuits and be admonished for doing so.[36]

A further process which enriches this side of our moral and emotional lives is the capacity for what has been called 'imaginative reconstruction'.[37] This may involve reconstructing in our minds the distress someone would have felt given their circumstances. It could also involve imagining how someone who is hitherto unaware of some disaster is likely to react on hearing the news, or even imagining how future generations will suffer as a result of some virulent disease or environmental disaster. In each case, the imaginative process can move us to a compassionate response, and thereby bind us to a much wider group than is possible when these responses first emerge.

I have already argued for the stable role emotions can play in our lives, and this is true of compassion also – the compassionate person is someone who is disposed to react and act compassionately on a consistent basis. This will receive much wider discussion in the chapter on moral reasoning, but we are now in a position to draw on the material in this chapter to provide a fuller description of the moral agent, where 'moral' is taken to indicate a compassionate perspective.

In the first instance, the moral agent is someone who has been brought up in such a way that she has acquired the right sort of emotional responses, a point made most famously by Aristotle.[38] This places a considerable degree of responsibility for our moral character outside our personal control, in that we are dependent on this around us for the right sort example when we start out. But this claim does not rule out personal responsibility for our actions, as reason also plays a central role in the life of the moral agent, and even where we can trace a generally non-compassionate set of dispositions back to a difficult background, we still expect the agent to be able to use reason to determine that actions about which she personally may be indifferent cross moral boundaries in ways which are considered to be unacceptable. Reason can provide a route into the moral perspective which can replace that of the ideal upbringing.[39]

---

[36] Raimond Gaita, *Good and Evil: An Absolute Conception*, 120.
[37] For discussion of this see Lawrence Blum, 'Compassion', in *Moral Perception and Particularity*.
[38] See in particular, *Nicomachean Ethics*, 1103b24-25.
[39] This will be discussed at much greater length in the next chapter.

The next key characteristic of the ideal moral agent is that she responds directly to the suffering of others. The internalised other of the moral agent includes a background of directed feeling which causes her to perceive suffering in a particular way. Such agents are distressed on coming to believe that someone else is suffering, and independently of concerns that they may suffer the same fate, or of thoughts that they would not like it to happen to them. It is this directness of this response and the corresponding lack of self-interest which characterise the moral outlook.

A third key point is that the mature moral agent, one whose moral perspective is a compassionate one, will bring to moral situations a degree of discernment. Less sophisticated compassionate responses, what I have called proto-compassion, involve an inability to distinguish between suffering which is important and is therefore to be alleviated where possible, and that which is not an appropriate object of concern. What is to count as important suffering will reflect a complex range of shared ideals and standards across a community. More straightforward cases will come in the form of sources of suffering which affect more or less any agent to a significant degree, such as physical pain or the loss of someone close. More complex cases will be those whose entry into our scheme of what counts as important can be explained less easily by means of our 'natural constitution'. If we take cases such as the loss of a centuries old family estate by an irresponsible heir, then one might plausibly argue that at least the suffering of other family members might be a form of suffering which merits compassion, or else one which is trivial in light of the fact that they didn't deserve such wealth in the first place.

The first and third points highlight in different ways the role of reason in the overall thesis I am presenting. I have claimed that reason can both lead us into the moral perspective, and that it is crucial once we are there. It is for these reasons that I have claimed that compassion and reason are jointly necessary and sufficient for a proper morality, and having provided a justification of the claim that compassion is necessary I will go on to justify the claim that reason is also necessary in the final chapter. But prior to this I wish to address the related issue of how one might use this approach to explain another area of moral thought which traditionally poses difficulties for those who wish to defend sentimentalism. How can a theory which claims that morality is grounded in the contingencies of our emotions explain practical necessity which morality involves?

Practical Necessity

In perhaps the most moving scene in *Anna Karenina*, Karenin returns home to Anna on receiving news that she is likely to die after giving birth to Vronsky's child. Having initially hoped for her death in light of the terrible humiliation he has endured due to her very public affair, Karenin is implored by Anna to forgive both her and Vronsky for what they have done. Faced with the depth of Anna's despair and the force of her appeal from what he takes to be her deathbed, Karenin finds it in himself not only to forgive them, but to offer comfort to Vronsky. 'Karenin took Vronsky's hands and moved them away from his face, terrible with its look of suffering and shame.'[1] But even more extraordinarily, he then sets aside consideration of the shame he himself has earlier suffered. 'You may trample me in the mud, make me the laughing-stock of the world, - I will not forsake her and will never utter a word of reproach to you.....I must and will remain with her.'[2] Vronksy is simply bewildered in the face of such a gesture, but not to the extent that he is wholly oblivious to the nature of what he has witnessed. 'He could not understand Karenin, but felt there was something high and inaccessible to one with his outlook on life.'[3]

When Karenin explains his actions, he tells Vronsky that 'the joy of forgiving has revealed my duty to me.'[4] And it is once the demands of duty have been revealed that it becomes clear to Karenin what he *must* do. And it is here where we begin to uncover an area of moral thought which raises a series of issues which sentimentalism must be able to explain if it is to be a plausible moral theory. Vronksy has not acted in a boorish manner, and has not set out at any point with the prime purpose of humiliating Karenin, but has acted primarily in pursuit of his love for Anna, who has acted with a corresponding desire to be with the man she genuinely loves. But it is Karenin who rises above the pull of inclination to act in accordance with what he believes to be right, as opposed to what he desires. From a Kantian perspective, there is an account of what explains Karenin's behaviour, and why neither of the other two protagonists are capable of the same sort of action. The 'high and inaccessible' point of which Vronsky is only vaguely aware is the moral law which commands us all insofar as we are capable of practical reason. Regardless of inclination, and not because of it that this law applies, and

---

[1] Leo Tolstoy, *Anna Karenina*, 412.
[2] *Ibid*, 413.
[3] *Ibid*
[4] *Ibid*

this law is the source of our obligations to one another. It is the recognition of his obligation as a Christian and a husband which grounds his *having to* remain with the wife he believes to by dying. It is because Vronsky has been guided primarily by inclination rather than a sense of duty, that he is incapable of comprehending Karenin's response to him, and it is why both he and Anna have conducted an affair which both of them believe to be inappropriate in light of Anna's marriage. If sentimentalism cannot explain the necessity which springs from obligation, then it fails to account of a key area of moral thought. It is therefore essential to provide such an account, and this will be the prime aim of the current chapter.

I shall begin by setting out Hume's account of the source of obligation (one to which I am broadly loyal), and place it in the context of the theory I have so far developed. I shall then consider one of the most powerful, recent Kantian critiques of this sort of theory, that proposed by Christine Korsgaard, and offer a response to her concerns for any moral theory which tries to explain the normativity of morality with reference to human nature. The overall position which I shall argue for is that one can provide an account of practical necessity which is consistent with a Humean approach, and which meets many but not all of the responses a Kantian is likely to raise. But I shall further suggest that Kantian concerns which cannot be answered need not be fatal to sentimentalism as it is unlikely that any theory could pass the kinds of test set out by Korsgaard.

Hume on Obligation

The key test for sentimentalism in this context is trying to provide an account of the law-like demand which obligation makes of us – the sense in which obligation compels us act in a particular way, often in opposition to our desires. There are moments when it is not simply that we feel we ought to do something, but also some reticence due to competing demands, but when we feel we *must* do it, and no competing demand carries any real force. What can explain the 'must' in such instances? Hume's account of obligation stems from his account of human nature.[5] Our basic constitution is one which combines 'selfishness and confin'd generosity'.[6] Our 'confin'd generosity' is further combined with

---

[5] Much of the following account follows that provided by David Wiggins in 'Categorical Requirements: Kant and Hume on Duty', in *The Monist*, Vol. 74, No. 1, 83-106.
[6] David Hume, *A Treatise of Human Nature*, 498-499.

our capacity of sympathy, which in Hume's terms is the tendency for persons to react in a similar to way to others one's they have recognised the suffering or the joy of the other person. This means that the kinds of actions which promote happiness in one person are likely to have a similar effect in others, and it is when such actions have a positive impact that they are called virtuous, and when they have a negative impact they are called vices. We very quickly come to recognise that unless we restrict our own selfish tendencies through the following of rules laid down by wider society, then we will not be able to enjoy the benefits of participating in our community. But whilst in the first instance we act largely out of concerns grounded in self-interest, we eventually come to realise that even very distant acts of injustice, which have no direct impact on our own situation 'are prejudicial to human society', and as a result 'we partake in their uneasiness by *sympathy*.'[7]

The key point which now emerges with regard to the question of obligation, is that we engage in a process of what is essentially moral induction, such that a 'general rule' can be derived from the limited number of instances we have observed. This rule covers not only our own actions, but also those of others who act in an unjust manner. In this way, Hume is able to explain the emergence of obligation out of a combination of different characteristics to be found in our natural make-up. '*Thus self-interest is the original motive to the* establishment *of justice: but a* sympathy *with public interest is the source of the* moral approbation *which attends that virtue.*'[8] Hume never claims that we lose our sense of self-interest, but simply that we come to recognise that observing the wider rules of society are actually the best form of satisfying our selfish desires. This combines with our capacity for benevolence to create a strong commitment to the obeying what we take to be the requirements of justice, and the condemnation of those who fail to observe them.

The most we can ever achieve in terms of improving the general moral standards of mankind is to work within the confines of what nature has imposed on us. 'The utmost politicians can perform, is, to extend the natural sentiments beyond their original bounds; but still nature must furnish the materials, and give us some notion of moral distinctions.'[9] The limited potential of what can be achieved in the public realm is mirrored in the home, where 'publick praise and blame encrease our esteem for justice;

---

[7] *Ibid.* 499.
[8] *Ibid.* 499-500.
[9] *Ibid.* 500.

so private education and contribute to the same effect.'[10] What we can hope to achieve in raising children appropriately, is that 'the sentiments of honour may take root in their tender minds, and acquire such firmness and solidity, that they may fall little short of those principles which are the most essential to our natures, and the most deeply radicated in our internal constitution.'[11]

To what extent then, could I claim support from Hume for the arguments I have made? Firstly, I suggest there is an obvious correlation between Hume's contrast between our selfish and generous impulses on the one hand, and what I have called selfish desires and the claims of the internalised other. Like Hume, I am committed to the view that these twin forces both exert a pull on us, and both provide us with reasons for action. Secondly, the idea that the emotions can be educated is one which I have suggested is central to both our understanding of the nature of emotion and to how we come to be concerned with the plight of others.[12] A third key point is an emphasis on certain emotions as the key to understanding what makes morality possible. Hume talks variously of benevolence and natural sympathy, whereas I have placed the emphasis on compassion.[13] The point of overlap is that each of these emotions is concerned with the welfare of others, and therefore open up the potential for a moral theory which avoids the sort of reductivism which would come if we were to try to ground morality in an emotion such as fear for our survival.

It is now important to return to the more specific question of how far this sort of approach can adequately explain the force of moral obligation. To begin with, it will be helpful to specify what is to be meant by the 'moral' part of the obligation. Within the context of the thesis I have set out, this will be an obligation to respond to the claims of others, where those claims can be understood as the need to relieve or avoid important suffering. Important suffering is to be understood in a very broad sense as enduring an unpleasant experience or missing out on important opportunities which can reasonably be provided.[14] What then can be said to explain the force which allows us to describe this as an obligation? One quick and easy answer which needs to be set aside is that an expression such as 'I had to help her' can be understood as expressing a thought such as

---

[10] *Ibid.* 500.
[11] *Ibid.* 501.
[12] See chapter three for full discussion of this.
[13] For the key distinctions between these concepts, see chapter six above.

'the reasons for helping her were much stronger than those of not helping', or 'I felt very strongly that helping her was the thing to do'. The reason why such an approach won't do is that, as Bernard Williams points out, 'Necessity is not the same as decisiveness.'[15] In practice this becomes particularly important when we realise that either of the putative analyses of the first proposition could be followed by 'but I still didn't help her.' If we are to take seriously the 'had to' or the 'must' of such expressions, then we must make sense of them as expressions of the *necessity* of complying with obligation, rather than just expressing a sense of feeling that there was a powerful but not irresistible reason to perform a particular action.

In light of this, the force of the sentiments which are inculcated in us must be such that they provide not simply a reason to act, but one which compels us to act in the face of whatever opposing reasons we may be aware of. This is the position I shall go on to defend, but before doing so I wish to address the response which wishes to rule out this sort of approach altogether, as it is this sort of approach which best sets out the nature of the challenge which must be overcome in order to provide a satisfactory account of the nature of moral obligation.

The Kantian Critique

The problem which any Kantian is likely to find in the sort of Humean account I have set out is that it fails adequately to address the demands which moral obligations place upon us. In Kantian terms, all moral imperatives on this account will be hypothetical ones, in that the most one can hope to achieve is the claim that if we are to be compassionate we will help those who are suffering, and if we are compassionate then this will lead to a richer life both for us and those whom we strive to help. But if this argument is addressed to someone for whom compassion has no pull, and who disagrees that it will promote his own interests to be compassionate, then there is no further appeal we can make from within the moral realm. This means that all those for whom the force of such appeals carries no weight have no reason to act in accordance with what I have called a proper morality, one which is consistent with our nature as human beings. Only a categorical imperative can provide the law-like necessity which applies to

---

[14] See 108-109 and 115-120 above for further discussion of this.
[15] Bernard Williams, 'Practical Necessity', reprinted in his *Moral Luck*, 126.

all agents, irrespective of inclination, and only practical reason can provide this inescapable source of normativity within the moral realm.

In order to respond to this sort of attack, I shall begin by setting out in greater detail why the sentimentalist account of obligation set out earlier must fail the Kantian test, and then suggest how one might respond. One of the most prominent, recent works of Kantian moral philosophy provides not only a detailed contemporary Kantian theory of moral normativity, but also an account of why any moral theory which is grounded in human nature must fall short of the required standard. In *The Sources of Normativity*, Christine Korsgaard argues that in order for a theory to stand up, it must be able to answer what she calls the 'normative question'.[16] This question is 'why must I do it?', where the 'must' is one which springs from the demands of morality. There are three criteria which must be met in order for there to be a satisfactory response.[17] The first is that the answer is one which the agent herself must accept, and it is not sufficient that we judge from a third person perspective that she must act in a particular way. It must be an answer such that the agent herself feels that in brings her moral enquiry to an end with a clear answer that she must act. The second criterion is that of 'transparency', which is to say that the agent must be aware not only of what she must do, but why. Justification is a part of any adequate answer to a normative question because we must understand the nature of the moral theory under which we are acting in order to be sure we are not acting under the influence of a motive we do not wish to endorse. Finally, the answer 'must appeal, in a deep way, to our sense of who we are, to our sense of identity.'[18] Korsgaard emphasise that the extent of what morality can demand of us, even giving up our lives in some cases, means that it must mean that on occasion the alternative to death is even worse – 'not being ourselves any more'.[19] A satisfactory answer to the question will bring the questioning to the end. 'The unconditional answer must be one which makes it impossible, unnecessary or incoherent to ask why again.'[20]

---

[16] Christine Korsgaard, *The Sources of Normativity*, 16.
[17] See *Ibid*, 16-17.
[18] *Ibid*
[19] *Ibid*, 18. This rather curious criterion is one criticised by Thomas Nagel in his response in the latter sections of the book (see page 206). As he points out, it would rather cheapen the motive of an act of self-sacrifice to say that we made it because we didn't want to lose our identity. The explanation ought surely to stop earlier with the claim that we made the sacrifice because we care about the survival of others.
[20] *Ibid*, 33.

The question now becomes whether or not the kind of sentimentalism I have defended can meet the criteria which Korsgaard has set down, and she suggests certain ways in which such a theory could not. The problem which Korsgaard highlights is essentially that once we reach the end of moral explanation in this kind of theory, we can still ask significantly if we must act in accordance with what it suggests our obligations are. If the question can still be asked in such a way that a reasonable person might still demand an answer, then we have not reached the 'unconditional' point that an adequate moral theory must lead us to. Such a response this can be deduced from her criticisms of Hume and Bernard Williams.[21]

She argues that Hume defends his moral theory from two different standpoints, that of self-interest and that of the moral sense.[22] Put briefly, she understands Hume to be claiming that our sentiment of benevolence makes it possible for us to be concerned with the welfare of others, and self-interest reinforces this on the grounds that we derive personal pleasure from the pride which stems from being virtuous. She understands Williams to be arguing that 'if we find that a social world promoted the best life or at least a flourishing life for human beings, this would justify the values embodied in that social world.'[23] According to Korsgaard, the similarities between Hume and Williams expose the limitations to each of their theories. 'Where Hume establishes normativity by showing that morality is congruent with self-interest, Williams asserts that it would have to be established by congruence with human flourishing.' In each case, we could meaningfully ask whether or not we must act in accordance with an obligation based on these premises, and the fact that the question stands in need of an answer indicates the inadequacy of the theory. The claims I have made over the last two chapters have drawn heavily on Hume, and the same question therefore emerges. If an obligation is one which stems from a morality based in compassion, and this sort of morality is consistent with the kind of being I am as well as one which will promote a richer life for me, must I obey it? If this is a genuine question, then it raises genuine doubts, and if there are genuine doubts then we do not seem to have the certainty embodied in the *must* of our moral obligations. For this reason, there is a need for an account of how sentimentalism can account for the force of obligation in the face of the normative question.

---

[21] See *Ibid*, 49-78.
[22] *Ibid*, 55.
[23] *I bid*, 75. It is worth noting that Williams broadly accepts this characterisation of his views, although he claims that Korsgaard somewhat exaggerates his neo-Aristotelianism. See *Ibid* 213.

To begin with, let me make plain that I do not believe that I can provide an answer which will satisfy the criteria Korsgaard has set down as they stand, but I do believe it is possible to provide an adequate account of how sentiments provide the force of moral obligation without meeting these criteria. More specifically, I believe that we can give an explanation of why we feel we must perform certain actions which accounts for our sense of practical necessity in ways which do not require the unconditional answer demanded by the Kantian. In order to elucidate this, consider for a moment the comment Korsgaard offers at the end of her attack on Hume, Williams and Mill; 'The test of reflective endorsement is the test used by actual moral agents to establish the normativity of all their particular motives and inclinations.'[24] This approach to morality has been earlier characterised as follows; 'If we find upon reflecting upon the true moral theory that we still are inclined to endorse the claims that morality makes on us, then morality will be normative. I call this way of establishing normativity the 'reflective endorsement' method.'[25]

It is not clear to me how literally we are intended to interpret the claim that 'actual moral agents' use this test. I take this group to be something like ordinary persons, in which case we are faced with two options. The first is that Korsgaard believes that ordinary people reflect upon whether or not they ought to endorse the claims morality makes of them. As a general empirical observation, this seems obviously false, and surely not what she intended. Alternatively and more plausibly, the above passage is intended to mean that ordinary people engage in moral enquiry in which they reflect upon moral questions and sometimes conclude that they must act in a certain way. If the first option is obviously false, then it is the second option we must consider, and this means that we need to explain the force of obligation in the context in which it appears, which is to say in ordinary discourse, and not as a feature of advanced metaethical theory. In this way, we may be able to explain the force of obligation in the context in which it appears in ordinary moral practice. It is to this that I shall turn in the first instance, before dealing

---

[24] *Ibid.* 89.
[25] *Ibid.* 49-50.

with the obvious riposte that this will not satisfy the categorical requirements of the moral law.

To begin with, it will help to specify with greater precision the realm of moral obligation. I have already said that the object of such obligations will be alleviating or avoiding the suffering of others within what I have called a proper morality, and the law-like sense is a reference to the 'general rule' which is acquired as a part of our emotional education. But further precision is required here. Firstly, the 'must' or 'have to' of moral obligation does not exhaust the modalities involved. Bernard Williams' work in this areas[26] focuses primarily on the impossibility of certain actions, and we also need to consider what is to be meant by expressions such as 'I couldn't just walk by and not do anything'. Next, we need to rule out certain senses of the modal expressions used. When we say 'I had to help him', we are not referring in this context to any external force which compels us act, as in 'I had to help him because he was holding a gun to my head'. And when we say 'I couldn't just walk by', this is not to be understood in the same way as 'I couldn't climb to the top of the tower because my vertigo prevented me from doing so.' The way in which we must understand incapacities in this context is as what Williams calls 'incapacities of character'.[27] These are to be understood in contrast to physical or psychological incapacities, and a distinct set of reasons to explain our action or inaction can be given depending on what the source of such incapacities are. If I say, 'I couldn't just walk by', then this might be explained by a physical force such as a strong wind preventing my advancing any further. Or it might be a psychological state such as a compulsive curiosity to see how people react when others help them up. It might also be out of ignorance, such as in cases where we might add, 'I didn't realise I could get away with it'. But it might also be for reasons which we express in distinctly moral terms such as 'I realised it was the right thing to do, and had to stop and help', and this fourth category is the one which leads us to the question of character.[28]

---

[26] See in particular 'Practical Necessity', reprinted in his *Moral Luck*, 123-131, and 'Moral Incapacity', reprinted in his *Making Sense of Humanity*, 46-55.

[27] See *Moral Luck* 129, and *Making Sense of Humanity*, 47 for further discussion of the different sense of the modal expressions in this context. It strikes me that Williams discussion owes much to Aristotle's in Book III of the *Nicomachean Ethics*. For detailed analysis of that position, see Terence Irwin, 'Reason and Responsibility in Aristotle', in A E Rorty (ed.), *Essays on Aristotle's Ethics*.

[28] One point of similarity in all three cases is that when we say 'I couldn't do it', this is followed by not doing it.

The way in which character can be used to explain our feelings of incapacity or our sense of having to act in a particular way is because our dispositions[29] condition the way in which we perceive the events with which we are confronted.

> What I recognise, when I conclude in deliberation that I cannot do a certain thing, is a certain incapacity of mine. I may be able think of that course of action, but I cannot entertain it as serious option. Or I can consider it as an option, but not in the end choose it or do it.[30]

Williams provides further detail of how exactly this process works in a later article, and once again, I shall quote at length.

> All this comes about because the dispositions that are the ground of the agent's incapacities are focused on to a particular case through the ways in which the feature of the case impress themselves on the agent – ways that are best represented by a deliberation, even though the deliberation need not consciously occur. It is because of this, too, that when there is a conscious deliberation, which issues in 'I can't', that conclusion presents itself at once as a decision, and as a discovery. It is a decision, as being indeed the conclusion of a deliberation whether to do that thing. But it presents itself to the agent also as a discovery, because the underlying dispositions have not before been focused through and on to that very conjunction of features.[31]

Within the context of what I have already set out on the subject of our emotional development, I would argue that the reason why certain features of a situation impress themselves on us in a particular fashion is a consequence of the background of directed feeling which provides the backdrop to our perception of the world.[32] We come to feel a particular way about certain features of the world, and part of that is constituted by our feelings concerning the needs of others. It is on perceiving someone in need that such feelings are brought into play, and it is this sort of response which allows us to recognise the dispositions both in others and in ourselves. As Williams points out, it is in part a

---

[29] The way in which our emotional development helps explain the development of dispositions of character is covered in 'The Moral Perspective'. See 76-87 above.
[30] Bernard Williams, *Moral Luck*, 128.
[31] Bernard Williams, *Shame and Necessity*, 52..

discovery when we realise that we cannot do something, and such discoveries can provide us with a better understanding of our own dispositions.

When we reach a point where we feel we must act in a particular way, or else we are incapable of acting, the necessity or incapacity indicates that the reasons for action or inaction are overwhelming. It is not that there need literally be no alternative, but rather that the alternative exerts no serious force in the range of options before us. We perceive a situation, and we are struck so powerfully by certain features of it that some actions will present themselves as either required or out of the question, and it is the dispositions of character to which Williams refers which explain why our perception of the situation operates in the way it does. As an extreme example of this, one might take the one offered by Hume in the second enquiry of the parent who 'flies to the relief of his child; transported by that natural sympathy which actuates him.'[33]

The reason why the need for action can press in on us so powerfully is because dispositions of character – kindness, concern for others, love of one's children, loyalty, etc - involve feeling a particular way about the world. When we are confronted with situations which bring those feelings into play, our response will necessarily be one involving our emotions, and emotions always have the potential to motivate us to act.[34] In some cases these situations will contain a level of urgency or extremity, or else such an unimportant series of alternatives, that one option will appear so crucial that its appeal overwhelms the motivational force of the others. In such instances, the motivational force of that option is such that it is the only one which exerts what we feel to be any serious pull, and it is in such cases that we feel we *must* act in a particular way.[35]

There is one important caveat here which Williams introduces. We might well be asked on occasion to betray a close friend, and find it impossible to do so, but it is entirely

---

[32] See chapter two above for full discussion of this.
[33] David Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 303.
[34] See chapter three for an account of the motivational force of our emotions. I say the have the *potential* to motivate us to act, on the grounds that there may be others motivations which are more powerful, and prevent our acting out of emotion on a particular occasion.
[35] John McDowell makes a similar point in 'Virtue and Reason', when he claims that our reason for action in such cases is seen, 'not as outweighing or overriding any reason for acting in other ways, which would otherwise be constituted by other aspects of the situation (the present danger, say), but as silencing them. Here and now the risk to life and limb is not even seen as any reason for removing himself.' Quoted from *Mind, Value and Reality*, 56.

possible that many of us will do so nevertheless under certain circumstances. There are forms of brutality or 'ingenious coercion'[36] which may result in our committing such acts of betrayal. But as Aristotle points out, 'In some cases, not praise but pardon is given, when a person does wrong because of things that strain human nature to breaking-point and no-one would endure.'[37] In such circumstances we might reasonably claim that this does not reflect the character of the agent, and that when we talk of moral incapacity we can do so meaningfully whilst acknowledging that dispositions of character are not invulnerable to other influences under all circumstances. I have already said that in order to understand what we mean by the term 'obligation' as it is used by 'actual moral agents', then we must consider how such terms apply in ordinary discourse. In light of this, it is perfectly consistent to claim that the agent who couldn't betray her friend under ordinary circumstances might do so under torture.[38]

In short, moral obligation in the form of either having to or being incapable of performing certain actions can be understood as a consequence of how we perceive the world in light of dispositions of character. In some cases, we find ourselves confronted by sets of circumstances which bring down certain emotional responses, because we are the sorts of people who are disposed to respond in this way when confronted with just such circumstances. In some of these cases, our perception of the situation will be such that only one course of action has any force in terms of what we are motivated to do or not to do, and this can be explained in terms of dispositions such as being compassionate or kind. The compassionate agent will perceive the suffering of someone in need as something which she needs to respond to, and on occasion the force of that need can explain why she can honestly report that 'I just had to help. There was nothing else I could do.'[39]

---

[36] *Making Sense of Humanity*, 54.

[37] Arisotle, *Nicomachean Ethics*, 1109a.

[38] Perhaps some qualification is required here. We take loyalty to be a characteristic we could not attribute to someone who consistently betrays a friend, but it would be a mistake to assume that this requires us to avoid any breaches of trust. All of us are prone to slip into disloyal attacks on friends at some stage, sometimes for the most trivial of reasons. Nevertheless, we are usually prepared to allow such a slip and to continue to think of the person who is guilty of it as a loyal friend. It is precisely for these sorts of scenarios that we have the expression to act 'out of character'.

[39] There is a need here to consider that contrast between statements such as 'I had to help' and 'I felt I should, so I did it even though I didn't have to'. This will be dealt with in the course of the next chapter on moral reasoning, as it concerns how it is that the process of deliberation can result in conclusions which have overwhelming force, as in the first example here, or ones which are less compelling, as in the second example.

In this way, we can provide an account of the force of obligation as it is used in ordinary discourse, but it is unlikely that this will satisfy a Kantian critic. There are two points at which such a theory is still likely to fall short of what Kant demands. Firstly, one could argue that the compassionate agent has not reached the unconditional response when she says she must do the compassionate thing, because she might still ask herself whether or not she really must do it. The second point is that if we are to ground moral obligation in the contingencies of human nature, then how are we to explain the intuition that even if people do not have any compassionate dispositions towards others, they must help others when they are suffering? Doesn't the obligation stretch beyond our contingent dispositions?

As far as the first of these criticisms is concerned, it is far from clear that this presents a genuine problem for the sentimentalist. The concern is that the compassionate agent may perceive someone in need of help, feel that she must help, but then ask herself whether she really must help because it is the compassionate thing to do. Once the question is asked in a serious way, then the doubt that it allows in diminishes the force of the original claim, and it therefore no longer appears to be an obligation. In order to draw the teeth of this challenge, we need to distinguish the different ways in which such an agent might ask this question of herself. The first way in which she might ask it is in the sense of whether or not there are possibilities which exist in the sense of their being physically or psychologically possible. The sense of the question in such cases would be similar to my asking if I could run the hundred metres in under ten seconds, or if I could stand firm in the face of battle in the face of overwhelming odds. Such questions require of an agent answers which indicate the range of possibilities from the third person perspective, and the agent can recognise the existence of such possibilities without acknowledging that they are possibilities for her.

The other relevant way in which the question might be a significant one for the compassionate agent is if she asks herself whether coming to the aid of someone in need, where there is no obvious hurdle to her doing so, is really something she must do, or if she couldn't simply ignore such a claim. If such a question really could present itself, and it is one that agent poses and understands in a distinctly moral light, then it would appear that we have not yet an adequate account of the necessity of moral obligation. For the obvious inference would be that if she genuinely doubts that she must do it, the

requirement of acting compassionately cannot be one she feels she *must* perform. We would need a weaker claim such as 'she felt very strongly that this was the right act'. But the simple answer to this point is that genuine doubt over whether or not to come to the aid of someone in need in such circumstances is not a possible thought for the compassionate agent. To be of a compassionate disposition means that she will perceive suffering in a particular light – as something to be ended or prevented as a matter of priority. Such a disposition is not simply one among many in the sense of competing with other dispositions as to what will determine our course of action when confronted with suffering. To describe it as disposition of character indicates that it plays a dominant role in our perception of situations where suffering is involved. Once we come to see suffering as something which requires our attention but in competition with non-moral demands such as the cost to us in time and effort, then the agent who has these thoughts just isn't one whom we would describe as having a deeply held compassionate disposition. As such, the normative question is not one which the compassionate agent can ask in the sort of way which would lead her to question the moral force of her obligation, and open up the need for a deeper foundation for her convictions.

The suggestion here is that a compassionate disposition of character blocks the potentially destructive path of the normative question. The answer to the question will be compelling for that agent whose dispositions condition her perception of the world in such a way that it seems obvious that suffering is a bad thing. Once the intellectual onslaught that Korsgaard wishes us to engage in is encountered, then the only way in which it will generate a diminution in the force that the disposition provides is if it first damages the disposition itself. We might become intrigued by the intellectual exercise of striving to suspend our tendency to respond compassionately to those in need, but this is already the thought of an agent whose dispositions have altered. To be sure, such dispositions are indeed vulnerable, and they may come to be worn down by misfortune or extreme circumstances, but this is a separate concern. There is as yet no reason to think that the compassionate disposition cannot stand firm in the face of the normative question, when that question is asked and answered in a distinctly moral light. As an example of this, consider once more Hume's example of the parent who flies to his child's relief. When asked later if he felt this was something he really had to do, then the genuinely loving parent could only answer in one way. The dispositions of character

which condition his perception of his child's circumstances make no other answer possible. If he were to say that he saw that helping his child was something he now realises he could have not bothered with, our horror would be compounded with the view that the parent lacks the dispositions which are required to describe him as loving of his children.

This still leaves in place the wider issue of how our obligations to our fellow humans rest on the contingencies of human nature, and this brings us into conflict with the thought that there are acts which would be wrong even if no-one objected to them. It is a perennial problem for any sentimentalist as to how to address Hume's 'sensible knave',[40] someone who has correctly calculated that is to his advantage to act unjustly, who feels no qualms about doing so, and whom we cannot bring to see that it is in his interests to act justly in this instance. The theory I have defended leaves open the possibility that if all persons thought as such, then no-one would have a reason to act justly because none of us would have the right sort of sentiments. But the thought stubbornly remains in place that even if none of us thought it right to treat others with respect, it would still be right to do so and we would all have reason to do so despite our inclinations. The only course open to sentimentalism here is simply to deny this is the case, and the best way to account for why we are so committed to the kinds of obligations we currently feel is best expressed by Hume.

> The interest on which justice is founded is the greatest imaginable and extends to all times and places. It cannot possibly be served by any other invention. It is obvious and discovers itself on the very first formation of society. All these causes render the rule of justice steadfast and immutable; at least, as immutable as human nature. And if they were founded on human instincts, could they have any greater stability?[41]

The personal needs, interests and characteristics which make justice both possible and necessary are amongst our very deepest. Our need for affection, our fear of suffering, our capacity for love and compassion, mean that in any community we are likely to see rules emerge which come to reflect in some form such primal features of who we are.

---

[40] See David Hume, *An Enquiry Concerning the Principles of Moals*, IX.2.
[41] David Hume, *A Treatise of Human Nature*, 620.

Our obligations are therefore ultimately contingent, but far from random. They are grounded in the deepest aspects of what it is to be human, and our needs and capacities are sufficiently stable to provide a robust basis for a powerful sense of obligation. As far as the concern over the ultimate contingency is concerned, we can do no more than observe that if the fundamentals of human nature were very different, then it is reasonable to claim that what we take our obligations to be might also be very different.

I do not expect these arguments to satisfy a Kantian critic. Korsgaard's demand for the unconditional response which will appeal on the basis of practical reason alone means that any space for a contingent basis for our sense of obligation must be ruled out. But I do hope to have provided an account of just how sentimentalism can explain the force of obligation without recourse to a moral law which holds independently of inclination.

8. Moral Reasoning

Introduction

I have argued that our emotions play a crucial role in our moral lives, and this brings us face to face with the question of the role of reason within such a theory. The need here is particularly pressing for the simple reason that emotion and reason are often taken to be antithetical to one another. In an autobiographical aside, Antonio Damasio states what is surely a common view of the relationship between these two aspects of our lives; 'I had grown up accustomed to thinking that the mechanisms of reason existed in a separate province of the mind, where emotion should not be allowed to intrude.'[1] Within a theory of ethics, such a state of affairs would raise a serious problem. We appear to reason in the ethical domain in a similar way to many other areas of practical reason - drawing inferences, assessing the consistency of our beliefs, coming to conclusions about the best course of action. But if reason and emotion are so distinct, then how could such moral reasoning take place at all? Wouldn't the emotions deliver up simple unreflective responses which guide our behaviour in the way fear prompts us to flee from danger?

Beyond explaining how the process of moral reasoning is possible, a further problem arises with regard to the stability of our moral views. Kant's view of inclination (under which he subsumed emotion) captured a powerful intuition about the problem of relying on emotion within a theory of morality, '[I]nclinations vary; they grow with the indulgence we allow them, and they leave behind a greater void than the one they intended to fill.'[2] This view of emotions as a capricious and unreliable conforms to a common view of emotions, and such criticism led Kant to regard with suspicion even our more praiseworthy ones.

> Inclination, be it good-natured or otherwise, is blind and slavish; reason, when it is a question of morality, must not play the part of mere guardian of the inclinations, but, without regard to them, as pure practical reason it must care for its own interest to the exclusion of all else. Even the feeling of sympathy and warmhearted fellow-feeling, when preceding the consideration of what is duty and serving as a determining ground, is burdensome even to right-thinking persons, confusing their considered maxims and creating the wish to be free from them and subject only to law-giving reason.[3]

Such thinking captures claims which must be dispelled if the theory I have set out is to stand up, and I shall set out an account of moral reasoning which explains how the emotions can provide a

---

[1] Antonio R Damasio, *Descartes' Error*, xi.
[2] Immanuel Kant, *Critique of Practical Reason*, 124.
[3] Ibid 124-125.

perfectly satisfactory basis for such reasoning once we have before us a more rounded ontology of the emotions than the slavish opponents of reason Kant was so concerned about.

In order to address the key issues surrounding the role of emotion in our moral reasoning, I shall look at four central questions, the answers to which will provide a general account of the nature of this process. In the first instance, I shall look at the question of whether or not there really is an opposition between having an emotion and good practical reason. After this, I shall look at the nature of personal deliberation, and at what we are doing when we engage in this. In particular I shall look at the issue of deliberating about ends and means, and discuss whether or not we can usefully talk of moral reasoning being concerned principally with one rather than the other. Once I have considered the nature of individual deliberation, I shall look at the nature of reasoning with others, and look once more about how and why this occurs. Finally, I shall make some brief comments about what conclusions moral reasoning can hope to deliver up, and in light of this, the implications at the metaethical level.

It will help at this stage to give an indication of the position I shall be arguing for. In the first instance, I shall show that the emotions play a far more important role in practical reason than the Kantian view acknowledges. I shall draw on the influential work of Antonio Damasio to suggest an account of practical reason in which emotion plays an essential role in the overall process of reasoning and acting on our conclusions. Far from disrupting the process, I shall argue that emotions provide the initial impetus for practical reason, then impel us through our deliberation, finally delivering up the required motivation to ensure action in line with the conclusions we reach. With regard to the process of individual deliberation, I will try to show that conceiving of this process as one of strictly ends and means presents a misleading picture, and that this should be seen as more of a heuristic process in which reasoning and acting come to alter what we see as appropriate ends and legitimate means to achieve them.

I shall also try to show that the process of individual deliberation is analogous to the process we go through when we reason with others. In both these processes, we need to explain how reasoning can take place at all, why we reason, and what we are trying to achieve. My aim is to show that when we reason individually we aim at right action, and when we reason with others we aim at bringing them to see things aright. Finally, I shall suggest such views reveal a series of sceptical concerns about the limits of moral enquiry.

The line of thought I shall be taking with regard to what seem initially like a series of distinct questions is unified by the overarching view that all of our moral reasoning takes place from within

what John McDowell describes as 'the midst of the way of thinking one is reflecting about'.[4] The image McDowell borrows from Neurath is of the sailor overhauling his ship whilst still at sea, and the central claim of this chapter is that this image captures the nature of our moral reasoning. If this is the case, and it is to be compatible with the theory set out over the previous chapters of this thesis, then it also requires that our emotions be conformable to such a process, and it is to this which I shall now turn.

Emotion and Reason

Before moving on to discuss the supposed opposition between emotion and reason, it will be as well to clarify the use of the terms in play here. I shall treat moral reasoning as a species of practical reason, which stands in contrast to theoretical reason. The principal contrast lies in the purpose of each, with theoretical reason aiming at describing the way the world is, and practical reason aiming at coming to a decision as to what action to take. To say that moral reasoning is a species of practical reason means that it is concerned with deciding what the morally right action would be in the circumstances. Although all moral reasoning will therefore be an example of practical reason, the latter has a much wider scope. It will include reasoning about quintessentially moral dilemmas such as whether or not to kill an innocent person to save other innocent lives, but also non-moral dilemmas such as deciding which exhibition to see given the aesthetic appeal of the various candidates.

In this section, I shall be defending the claim that emotion is essential for effective practical reason, but let me begin by admitting the obvious fact that there are clearly times when our emotions disrupt the smooth progress of reason. The most obvious examples are in the kinds of cases Peter Goldie is referring to when he talks of 'when the red mist comes down over the eyes, and we can feel the blood pulsing in the temples, things look other than the way they are, and, accordingly, our emotions can mislead us profoundly'.[5] Such disruption can occur in a number of ways. We can be so jealous of someone that we cannot treat them with the fairness we believe they are entitled to. We might be so distracted by our smouldering anger that it becomes impossible to focus on a task at hand. Or we might be so crushed by a sense of grief that no issue seems worth thinking through in order to reach a decision as to what to do about it. In each case our emotions disrupt or block the process of practical reason, thereby preventing us from reaching what we ourselves would consider the better decision we would have come to were we not in the grip of these emotions.

---

[4] John McDowell, *Mind and World*, 81.
[5] Peter Goldie, 'Can We Trust Our Emotions?', in *Richmond Journal of Philosophy*, Vol, Issue 1, Summer 2002, 27.

To admit that there are some cases where emotion prevents good reasoning is clearly insufficient for the conclusion that in all cases emotion prevents good reasoning. Indeed, notwithstanding such examples, I shall suggest that all cases of effective practical reason require our emotional engagement, and that when our emotional lives are disrupted we cannot engage in such reasoning. When I talk of 'effective practical reason', I shall take it to be composed of three constituent parts. The first is an ability to think through all the information at hand, weighing the different factors appropriately. The second is an ability to come to a decision as to what to do in light of the prior process of deliberation. And the third is a capacity to act on the decision. What follows is a conceptual outline of how emotion aids us in the process of practical reason, and empirical evidence to support it.

Our emotions orient us towards the world.[6] Certain features of the external world have an emotional impact on us, whilst others leave us cold. When we are presented with certain situations which are frightening, disgusting, worrying, unpleasant, attractive or classifiable under a wide range of other such 'feeling-laden' descriptions, we may be prompted into immediate, unreflective action[7] such as in cases of 'blind panic', or we may be prompted into reflection about what to do about it. The latter cases bring us to the process of practical reason, but why are we prompted into it in the first place? The situations which provoke us provoke certain feelings, which I have characterised as 'perceptions of the state of one's body'.[8] Feelings have a variety of properties, including being pleasant or unpleasant. The phenomenon with which we associate the feeling, which could be an action, an object, a person or a state of affairs, either attracts or repels us as a result of the feelings we have about it.. We are motivated to reason what to do as a result of the pleasant or unpleasant nature of the feeling, and the feelings guide the ensuing process.

To put it more crudely, we either want more of what it is that attracts us, or less of what repels us. Of course this fails to capture the subtlety, complexity and depth of our considered moral reflection, but my aim here is to do nothing more than outline the properties of such reflection which are responsible for initiating it in the first place, driving it forward, and then producing action. I shall return to the danger of reduction in a moment, but it is worth emphasising that in very different ways both Hume and Aristotle also place great importance on the nature of our feelings within the process of morality. For Aristotle, 'it is pleasure that causes us to do base actions and pain that causes us to abstain from fine ones,'[9] and it is for this reason that it is crucial for us to have the right sorts of feelings if we are to be virtuous agents. For Hume, 'We do not infer a

---

[6] See chapter three above for further discussion of this.
[7] For a full discussion of action caused by emotion, see Peter Goldie, *The Emotions*, 37-47.
[8] See chapters two and three below for the full discussion of both what feelings are and how they come to be influenced.
[9] Aristotle, *Nicomachean Ethics*, Bk II, 1104b, 10-11.

character to be virtuous because it pleases: But in feeling that it pleases after a certain manner, we in effect feel that it is virtuous.'[10] Once again, the feelings we have will be of central importance in shaping our attitudes towards the world, and the way we reason in light of those attitudes. If this is right then we need to understand exactly how our feelings operate within that overall process.

In the first instance, the account set out here indicates that we are hurled into reflection rather than choosing or deciding which issues we will devote ourselves to. There is a passive element to our emotions and feelings, a sense in which they are already responding on our behalf before we come to reflect about the circumstances in which we find ourselves. We are already immersed in a world prior to any reflection, and this brings with it certain involuntary reactions to events which confront us. We see someone in pain, and if we are compassionate we react on seeing the person, prior to any process of reflection which then takes place, but in a way which provides the impulse for that process.

The process of deliberation, then, is impelled forward by feelings, but how does it operate? In a different context, and in the midst of a very different theory, Thomas Nagel sets out certain criteria to which any sound theory of practical reason must surely conform.

> Beliefs provide the material for theoretical reasoning, but finally there is something besides belief, namely reason, which underlies our inferences from one set of beliefs to another, and explains both the conclusions and those logical beliefs which embody our inferential principles in general propositional form.[11]

Such views look initially as if they represent a clear obstacle to the kind of non-cognitivist account I am arguing for, in that I have already accepted that our emotional episodes are capable of disrupting the inferential passage of one belief to another. But this obstacle is illusory in light of the way I have characterised the intentional and complex character of our emotions and feelings. I have referred already to the term 'feeling-laden belief' deployed by Michael Stocker.[12] Such beliefs differ from their neutral counterparts in the sense that they provide a different dimension to the process. If I am deliberating between watching a soap opera on television and going to the funeral of a close friend, then the latter seems to a exert greater pull than the former. The reason it is distinct is because concepts such as friendship entail having certain feelings about the person we see as a friend. Once those feelings are present, then they guide our reflection, providing it with a distinct content. Going to the funeral now becomes an act of loyalty, because it involves answering the

---

[10] David Hume, *A Treatise of Human Nature*, 471.
[11] Thomas Nagel, *The Possibility of Altruism*, 31.
[12] See page 28 above.

claim of friendship, a claim which is *felt*. One just could not sincerely consider someone to be a friend if one were indifferent to her, and it is here where the contrast comes in with the watching television. Such an option may well offer an appeal which is also felt, but it does not yet bring into play feelings about others, and does not therefore bring into the process of reflection the as powerful a set of loaded, feeling-laden concepts such as friendship or loyalty. The presence of certain feelings which are intricately bound up with the beliefs dramatically alter the path the reasoning takes. Although it scarcely does justice to the role such beliefs can have, the term which captures the way in which such beliefs contribute to the process is that they carry greater *weight*, and as such, this category of belief has the ability to overwhelm the weaker counterparts.

It is open to the Kantian at this point to counter that the claim of friendship need not be conceived of in terms of feelings. Could it not be the call of duty and respect for the moral law which explains why it is that such claims weigh more heavily with us? This strikes me as a plausible suggestion, but my principal focus at this stage is to set out a coherent account of how emotion and feeling can be reconciled with the process of practical reason. When considering the empirical evidence, I shall go on to say that this offers more support for a sentimentalist account of practical reason than a Kantian one.

To return to the point about the role of emotion in moral reasoning, a sceptic could simply accept the premise about motivation, in that it would be difficult to deny that emotions certainly motivate us to act in certain ways, and as we have seen, Kant will certainly grant the power of certain 'good-natured' inclinations. But the question still remains as to whether or not the 'feeling-laden' beliefs I have referred to could conform to the standards of reason which Nagel referred to. If our moral reasoning involves feelings and emotions, then don't they invariably bring with them a non-rational element which will get in the way of drawing inferences and reaching conclusions which conform to the cold hand of reason?

I acknowledged at the beginning of this chapter that when the 'red mist' descends it is unlikely that we will be able to reason effectively, but I have also stressed throughout this thesis that it is a mistaken view of our emotional lives to see it entirely in terms of unreasoned outbursts. Our emotions are often stable, open to change based on reasoned argument, and entirely compatible with a central role in practical reason. In order to highlight this, we can consider they way in which we expect coherence and consistency amongst our own views and those of others when it comes to arguing in ways which are clearly infused with emotional responses to a state of affairs in the world around us. Consider the following example:

It is morally wrong under all circumstances to support a law which permits the lives of the innocent to be taken.

In countries where the death penalty operates innocent people are executed.

Therefore

Supporting a law permitting the death penalty is morally wrong.

Even if the person offering this argument did so using the premise that emotion was the basis of all moral commitment, this wouldn't lead us to doubt that this is a sound use of the practical syllogism. Once we have a fuller understanding of the stability, persistence, and coherence of our feelings, and the way in which they come to colour our beliefs in a consistent manner, then we have taken a major step in explaining why a process of reason which has a central role for emotion leaves open the possibility of rational criticism and rational deliberation. The reason why this is so is because the way in which we acquire the direction of feeling involves the consistency that sound practical reasoning requires.[13] We strike someone, and we receive a look and tone of admonishment. We do the same thing again with the same result, and we come to associate the act of violence with the stern response, and thereby come to feel a particular way about violence. There is sufficient coherence built in to the education of the feelings and emotions, an education that entails the combining in consciousness of the unpleasant feeling and the belief that we have struck someone, that feeling and emotion can conform to practical reason in view of their association with certain relevant beliefs. The beliefs provide, as Nagel suggests, the material for the process of reasoning, but we are dealing with a material of a distinct nature, distinct in light of the feelings which they now bear.

The best way to capture this process within the current discussion, is to say that our feelings towards the world make certain aspects of the world *salient* when we come to consider which course of action to take. When we have developed an aversion to violence, the perception of violence now becomes salient in our process of deliberation in ways which would not have been possible without this sort of emotional education. This phenomenon of certain states or activities coming to be perceived, so to speak, under the influence of feeling, is what it is for such states or activities to take on greater weight within our process of reasoning.

In short, feelings impel us into the process of practical reason in virtue of the fact that they just are a response to what is perceived in the world when they are experienced within emotions, emotions

being intentional states. We have already been drawn in before we come to reflect on what to do, in that we already have pro or con attitudes with regard to certain features of our situation.[14] In addition, they provide greater weight for beliefs about some courses of action rather than others, and the fact that there is a level of consistency and coherence in the beliefs they colour means they present us with the possibility of rational criticism of our practical judgements. That is to say, a con attitude towards violence will emerge across a range of different circumstances and scenarios which involve actual or possible violent acts. Where there is such consistency, then there is the possibility of criticism whenever we find inconsistency in our responses, or lack of coherence in our beliefs with regard to violent acts.

So our feelings provide us with the impetus which drives us to reaching decisions. It is only once a decision has been reached that appropriate action can commence to resolve the problem, and it is for this reason that feelings push us through the process of deliberation until a satisfactory outcome has been settled upon. And this structure is reflected in our manner of explaining our actions and those of people around us. We see someone being attacked. The belief about the violence of the attack is combined in consciousness with a feeling. This combination of these two elements is an emotional response, which might be, in this case, anger at the aggressor. This means we have already responded before we come to reflect on what to do, and that response is the emotion. The anger involves a con attitude, which explains why we contemplate action at all, and if we decide that the thing to do is to intervene, the anger will then offer an explanation at two levels. If asked 'Why did you even think about doing something?', and then 'Why did you do something?', the answer 'Because I was angry.' rationalises both our entering into the process of practical reason and the ensuing action, which is to say the anger is a reason for both, it explains both, and it caused both.[15]

The third and final element of effective practical reason is that appropriate action must emanate from it. It is not sufficient that we come to a view as to what action to take - we must also take that action, and we must take it because of the decision we have come to.[16] As such, feelings spur us into reflection in the first place, then impel us through the deliberative process, and then provide the impetus for action. But beyond this, there is a further role they must play. The decision which

---

[13] See chapter three above for the full discussion of how our feelings come to be directed.
[14] For fuller discussion of the relationship between feelings and pro and con attitudes, see page 39 above.
[15] See 39-42 above for a more detailed account of emotion and motivation. I am aware that this may seem like a painfully analytic consideration of an action whose phenomenology may be much more unified and swift than this suggests, but this is not intended as an account of the phenomenology of such angry responses. There will be a discussion of what I call 'simple moral responses' in the final section of this chapter.
[16] The position I argue for here is consistent with that expressed by Davidson in 'Actions, Reasons and Causes', reprinted in his *Essays on Actions and Events*. Like Davidson, I hold that reasons cause actions, and I would wish to argue, which Davidson does not, that the causal force is provided by our feelings and emotions. A further consequence of this thesis is therefore that all reasons for action are 'internal reasons', in the sense argued by Bernard Williams in 'Internal and External Reasons', reprinted in his collection *Moral Luck*. On my account of

issues from deliberation may well require an ongoing commitment on the part of the agent. A decision such as committing oneself to working in the public sector for the sake of one's fellow citizens requires a sustained engagement with certain activities, a disposition to ignore the appeal of alternatives, a readiness to persist in a particular career even during periods of particular dissatisfaction. It is what I have called a background of directed feeling[17] which makes this possible. When we reflect on which course of action to take, we arrive at a decision which involves coming to feel a certain way with regard to the action we have settled on. This guides us in a settled, stable, ongoing fashion, which persists despite the kind of obstacles I alluded to earlier. This is not to say that such feelings are static, and many will be open to revision or even rejection further down the line, but my point is simply that they provide the platform for an ongoing implementation of the decisions issuing from practical reason, one which complements the motivational force they offer at the moment of the decision being taken. Feelings perform this role in light of their making certain features of the world more salient within the deliberative process, often on the kind of ongoing basis required for sticking at certain projects.

Before setting out the evidence which supports this conceptual outline, it is worth emphasising one point with regard to the limitations of what I am arguing for. What I have set out is no more than the form of how emotion and feeling contribute to our practical reason, but once one acknowledges the crucial role for such phenomena, then there is a constant danger of appearing hopelessly reductive in the descriptions we offer. I have argued already that one problem we face in describing our feelings and emotions is the impoverished vocabulary at our disposal when it comes to saying what they are like, and it is crucial not to lose sight of this point when discussing the way we actually reason. There is depth of what it is like, of *how* it feels to love or hate, to grieve or to fear, and the argument I have set out here should be seen as a formal outline of the reflective process often prompted by such emotions, and nothing more. The phenomenal feel of such thinking will often be far better explored in poetry and literature, and by way of example consider one of the seminal cases of practical reason from literature:

> '"You mean I have to choose?"
>
> "You're a Polack, not a Yid. That gives you a privilege - a choice"
>
> Her thought processes dwindled, ceased. Then she felt her legs crumple. "I can't choose! I can't choose!" She began to scream. Oh, how she recalled her own screams! Tormented angels never screamed so loud above Hell's pandemonium. "*Ich kann nicht wahlen*" she screamed.....

---

reasons for action, a reason will come out as a combination of a belief and a feeling combined in consciousness, and which would have to form part of the agent's 'subjective motivational set'.

[17] See chapter three above.

> "Don't make me choose," she heard herself plead in a whisper, "I can't choose."
> "Send them both over there, then," the doctor said to the aide, "*nach links.*"
> "Mama!" She heard Eva's thin but soaring cry at the instant that she thrust the child away from her and rose from the concrete with a clumsy stumbling motion. "Take the baby!" she called out. "Take my little girl".[18]

To talk of feelings and beliefs as anything more than a rather lifeless, technical description of the reasoning which goes on here would be grotesque, and this point will be crucial when I return to the question of how we reason with others. But I cite this here merely to emphasise there is a need to be aware that discussion of our emotional lives through analysis of concepts or consideration of biological factors can be expanded to include the kind of exploration found in literature, which may sometimes provide us with a fuller account of what we are discussing. What I wish to consider now is whether or not there is wider evidence to support the theory as a whole.

Phineus Gage Matrix

If the outline set out above is correct, then it entails certain claims which can be measured against empirical evidence. One such claim is that if a background of directed feeling is essential for the smooth operation of practical reason, then in cases where someone is incapable of standard emotional and feeling-related responses, they should be incapable of effective practical reason. A second claim is that if we are capable of reasoning at an abstract level about what course of action one should take, this will be insufficient to generate action once a decision has been reached, or else insufficient to make the decision stick where one is required to act consistently over an extended period unless there is a feeling or emotion involved. Both of these claims are supported by the findings of Antonio Damasio's research into patients who have suffered brain damage to areas such as the ventromedial prefrontal cortices and the somatosensory cortices in the right hemisphere.[19] Damasio describes the condition he has identified as the 'Phineus Gage Matrix', after a nineteenth century American railroad worker who suffered but survived horrific injuries to the prefrontal lobes as a result of an explosive charge prematurely detonating.

Damasio analyses the medical history of Gage and similar victims of brain damage to these areas of the brain which are regarded as the seat of the emotions. With Gage, as with the others whom Damasio studied, bodily survival was accompanied by a transformation in character. 'Gage's disposition, his likes and dislikes, his dreams and aspirations are all to change. Gage's body may be

---

[18] William Styron, *Sophie's Choice*, 589-90.
[19] Antonio Damasio, *Descartes' Error*.

alive and well, but there is a new spirit animating it.'[20] Part and parcel of this transformation was an inability to interact successfully with those around him, as he had done before. An increasingly chaotic social life in which he was unable to sustain previously stable relationships, develop new ones, or remain in work, saw him finish up in a circus before an early death.

One of the patients Damasio was able to study in person provides further evidence for the view that disruption in the emotional life severely impairs our ability to reflect and act effectively. 'Elliot' had had a brain tumour successfully removed, but not before it had damaged the frontal lobes against which it had been pressing. The impact on Elliot was that he could no longer organise his life in the way he had done previously. Damasio describes his behaviour as follows;

> He needed prompting to get started in the morning and prepare to go to work. Once at work he was unable to manage his time properly; he could not be trusted with a schedule. When the job called for interrupting an activity and turning to another, he might persist nonetheless, seemingly losing sight of his main goal. Or he might interrupt the activity he had engaged, to turn to something he found more captivating at that particular moment.[21]

In short, Elliot had lost the capacity to plan, organise and prioritise either short-term or long-term tasks. But there is a further dimension to his condition of particular importance to the theory set out earlier. The psychological analyses of Elliot indicated that his intellectual abilities, far from being impaired, were either average or superior, including in tests where he was asked to judge which course of action would be appropriate in moral dilemmas, where his answers indicated an outstanding ability to think through respond in ways which were consistent with the responses of those who had suffered no such brain damage.

The area where Damasio concluded that Elliot was wildly at odds with others who participated in the tests he undertook was in his emotional responses. In tests where he was shown emotionally charged images such as people injured in terrible accidents he was unmoved, and even when describing the downward spiral of his own life he spoke as a 'dispassionate, uninvolved spectator'. What adds to both the tragedy and the curiosity of the case was that he was aware of how different this was to the way he would have responded in the face of such circumstances prior to the tumour taking hold. 'In some curious, unwitting protective way, he was not pained by his tragedy.'[22]

---

[20] Ibid. 7
[21] Ibid. 36.
[22] Ibid. 44.

To what extent does this provide support for the account of practical reason I set out earlier? Firstly, Damasio himself concludes that our capacity for feeling provides us with the 'somatic markers' which he describes more colloquially as 'gut feelings'. These attract or repel us from certain stimuli, and he sees them as essential to our capacity for impelling us to act on a decision, and to stick at it over time. Further support comes in the form of the ability of some of Damasio's patients to continue to reason through abstract problems in ways which do not translate into an ability to act effectively in real life. Elliot's performance in the laboratory tests was exemplary, yet he simply could not translate such ability into effective action. The reason given for this is his emotionally flat response to the situations he was asked to address. In the terms set out at the beginning of this chapter, it is as if conclusions reached at the end of a process of practical reason which would normally constitute reasons for action, were treated as conclusions of theoretical reason which did little more than describe the way the world is. They either failed to generate the actions they should have or else failed to provide the persistence which was required.

It is at this point that we find a particularly strong challenge to the Kantian view. If respect for the moral law provides the required motivational force for us to perform our duty, and conclusions reached about what to do require no additional motivational support from the emotions, then Elliot's condition appears difficult to explain. The evidence suggests that he retained an outstanding ability to reach conclusions about how to act, but clearly lacked the ability to act accordingly. One response might be to ask how confidently we can count on the empirical evidence given the handful of relevant patients Damasio has treated, and it would be over-confident to treat this as a knockdown argument, but it adds force to the view that the emotions are necessary for our practical reasoning given the detailed analysis of the individuals Damasio has looked at. It also conforms to a standard Humean claim that reason on its own cannot motivate us, and it coheres with the wider claims contained in the present outline of our process of practical reasoning.

The conclusions reached by Damasio are consistent with the line of argument developed in chapters two and three above, and at the beginning of this one. We are oriented towards the world by background of feeling which comes to be directed in a variety of ways over the course of our development. It is when we are confronted by unusual or pressing circumstances that these feelings often erupt within the context of our emotions, and it is only when these feelings are operating normally that we can deliberate in ways which will allow us to act on the decisions we reach.

Individual Deliberation

Having set out this outline of how feeling and emotion contribute to practical reason, it is now time to focus more precisely on the subject of moral reasoning. When someone deliberates about what

would the morally right thing to do would be, how should we understand this? Firstly, we need to have a clearer idea of how the moral part of the question distinguishes these examples of practical reason from others. I have characterised the moral perspective as an aspect of the internalised other[23] - the feelings and emotions which have been shaped by those around us such that this influence has become part of our own perspective on the world.

Moral deliberation will aim at deciding on the right action in light of the claims, needs and interests of the others around us. It will involve the same processes of reason as other forms of practical reason - deciding what the relevant considerations are, considering the probable outcome of different actions, drawing inferences from our premises, aiming for a conclusion which will constitute our reason for acting - but the content of this sort of practical reason will be distinct. It will stand in contrast to deciding what to do in order to satisfy what I have called personal desires. The involvement of other persons generates a series of feeling-laden beliefs which provide the content of our reflection with a significance beyond that of deliberation about more self-regarding subjects such as which television programme to watch.

How then should we understand the deliberative part of moral deliberation? One very straightforward way of breaking down the problem is to say that either Aristotle was right when he argued that is deliberation of always of means, or else Aurel Kolnai was right when he entitled one of his articles 'Deliberation is of Ends'.[24] This might be the most straightforward approach, but it is unlikely that it would provide a satisfactory theory in this context. Kolnai's title is deceptive in light of his discussion of the complex way in which ends and means relate to one another, and at least one important interpretation of Arisitotle's dictum suggest that we need to be on guard against being too quick in how we understand what Aristotle meant.[25] I shall nevertheless approach this part of the discussion by focussing on the question of means and ends, but argue that in the final analysis that they are interdependent.

Let us first consider the starting point of any instance of moral deliberation. The first crucial observation is that the theory of emotion set out so far indicates that there could be no such thing as a neutral standpoint. We arrive at any situation with a background of directed feeling which orients us towards the world around us, and which involves a way of seeing and understanding what we find before us. This orientation is acquired from a very early age, and develops as we do. This means that we are disposed to respond in certain ways which reflect a wide range of influences

---

[23] See chapter four above for a full discussion.
[24] See Aristotle, *Nicomachean Ethics*, Bk III, 1112b, and Aurel Kolnai, 'Deliberation is of Ends', *Proceedings of the Aristotelian Society* (1962), 195-218.
[25] David Wiggins, 'Deliberation and Practical Reason', in *Needs, Values, Truth*, 215-239.

endured over the course of a life.[26] This much may be admitted without necessarily impinging upon a debate about means and ends, in that two people from radically different background might still be expected to converge on the same answer as long as we could isolate a series of moral facts which any rational individual could recognise. But the theory I have argued for rules out such an option with regard to moral deliberation on the grounds that both means and ends are dependent not only on one another, but also on the perspective of the agent herself.

Let me offer an example from literature. Ian McEwan's *Atonement* tells of an adolescent fantasist, Briony, who is deeply suspicious of the man, Robbie, who is pursuing her elder sister. Her suspicion is such that she opens a letter from Robbie intended for her sister. She knows that 'It is wrong to open people's letters, but it was right, it was essential, for her to know everything'.[27] Her sincere conviction that 'it was right' stems from her genuine fears about Robbie, and her heartfelt desire to protect her sister. On reading the letter, an obscene note which has been mistakenly placed in the envelope, Briony is reinforced in her previous view of the situation. 'With the letter, something elemental, brutal, perhaps even criminal had been introduced, some principle of darkness, and .....she did not doubt that her sister was in some way threatened and would need her help.'[28] The man not only presented a fearful picture to Briony, but 'disgusted her profoundly.' By this stage she is convinced there is a genuine threat to her sister and the wider household, and she is motivated to protect both. She seizes her chance when another girl visiting the house is raped, and Briony dishonestly identifies Robbie as the assailant, resulting in his ultimate imprisonment for a crime he has not committed.

When reflecting, she is aware of certain doubts regarding her course of action, but remains constant nevertheless.

> [t]he glazed surface of conviction was not without its blemishes and hairline cracks. Whenever she was conscious of them, which was not often, she was driven back, with a little swooping sensation in her stomach, to the understanding that what she knew was not literally, or not only based on the visible. It was not simply her eyes that told her the truth. It was too dark for that. Even Lola's face at eighteen inches was an empty oval, and this figure was many feet away, and turned from her as it moved back around the clearing. But nor was this figure invisible, and its size and manner of moving were familiar to her. Her eyes confirmed the sum of all she knew and had recently experienced. .... The truth

---

[26] See chapter three above for a full discussion of this.
[27] Ian McEwan, *Atonement*, 113.
[28] Ibid, 113-114.

> instructed her eyes. So when she said, over and over again, I saw him, she meant it, and was perfectly honest, as well as passionate.[29]

I suggest Briony's behaviour is representative of the *process* of moral reasoning in general. In the first place, she starts out with a series of relevant feelings which orient her towards the particular circumstances she is confronted with - her suspicion of Robbie and her love of her sister. These feelings then shape her understanding of subsequent events. She judges that opening the letter is right because of the overwhelming need to protect her sister and the household from Robbie, and believes lying to the police is justifiable on the grounds that her understanding of the truth has been enhanced in a way which may escape them if the bare facts are presented. What she 'knows' extends beyond what she saw with perfect clarity, and this honest conviction provides her with the basis of her testimony. In the terms set out above, her background of directed feeling conditions her understanding of events and then guides her as to the right course of action. At each point, it is her emotional responses and her feeling-laden beliefs which elicit reactions with a particular content, and prompt her in certain directions. Her understanding of which actions would be right and wrong is conditioned by the subjective standpoint of her emotional make-up. She reasons from the premises that these responses provide her with to conclusions which guide her ensuing actions. The problem now lies in explaining how someone who is reasoning in what I have suggested is a representative way has got things so horribly wrong, and requires an explanation of what is needed for us to get it right.

Practical Wisdom

It is as well to begin by pre-empting the obvious objection that this example helps to demonstrate that once you allow the emotions to intrude into any process of reasoning then this sort of disaster is likely to ensue. I have argued that Briony's behaviour is representative of the process of moral reasoning, which is distinct from the claim that it is an instance of good moral reasoning, and to use the example to impugn any emotional involvement in reasoning would be like taking the mistakes of a child in her maths exam to be representative of the redundancy of mathematical reasoning. What is required is an explanation of what has gone wrong, and what has to happen for things to go right.

One of the most interesting features of Briony's case is that she starts out with many of what one might intuitively consider to be the right emotions - love of her sister, fear of threats to her family, an attachment to ensuring that things turn out as they should. She is genuinely motivated by a concern for others, and it is this which makes possible her conversion in the later sections of the

---

[29] Ibid. 168-169.

book which provide it with its title. Years later, when she has reflected and confronted what she has done, her 'familiar guilt' would 'pursue her with a novel vibrancy'. The transformation she undergoes leaves her with an intolerable sense of her own wrongdoing. 'She felt the memories, the needling details, like a rash on her skin.'[30] When, years later, she meets Robbie again, she is tempted to defend herself on the grounds that 'She hadn't intended to mislead, she hadn't acted out of malice.'[31]

I suggest there is an obvious connection between the decency of her motives when she was younger, and the motives for her atonement when she is older. It is because she has a concern for at least some others that she is capable of coming to see her own actions in a different light. The compassion and concern which are part of who she is offer the initial motivation for reflecting upon her actions and drive forward the process which culminates in remorse for her previous actions and the desire for atonement. But why did she go wrong at first?

One factor which seems central to answering this is Briony's extreme youth when she acts to incriminate Robbie, and it is surely central any correct explanation that her change of heart occurs over a period of years which see her pass from adolescence into adulthood. In order to understand the change and to move us closer to an idea of 'getting it right' when it comes to moral reasoning, we can distinguish between two different senses of getting it right. One possibility is that Briony has come to understand a series of true propositions which she did not know before, and this has altered her understanding of events. But this just seems wrong at two levels. Firstly, Briony has not come to be aware of any new facts. She is aware throughout that opening other people's letters is wrong, that lying is wrong, that she did not clearly see Robbie's face, that she did not state the bare facts of the matter to the police. In other words, the transformation which moves her from getting it wrong to getting it right does not involve learning new facts, but a different way of judging the old ones. The second sense in which this explanation seems wrong is at the purely intuitive level, in that the sense of guilt, the desire to atone, the constant shame, seem to be bound up in an essential way in the changes which have taken place, but seem out of place in any explanation which centres around learning some new facts. It is at a profoundly personal level that the key changes have taken place, and that very personal aspect points us surely in the direction of the emotions than towards the acquisition of factual knowledge.

To say that the key changes have taken place at a profoundly personal level means that Briony has come to 'see' events differently, that she has acquired different feelings about the actions of others, and about her own actions. Once the feelings change, this is a change in her attitudes. Events

---

[30] Ibid, 324.
[31] Ibid, 336.

which previously elicited one response now elicit a different one, and this new way of feeling about events drives her reasoning in a new direction, and generates new conclusions. McEwan provides us with frustratingly little material about the process of change in his character, but we can speculate as to what a good explanation might be. One key element would surely be a greater humility in the confidence Briony has in her own convictions, a greater awareness of her own limitations when it comes to a full understanding of events. There might also be a greater deference to the ability of others to reach sound judgements in the face of the facts, and a corresponding increase of faith in the effectiveness of existing procedures for reaching good decisions. Finally, there could be a heightened understanding of the awfulness of what she has done to Robbie, and the unreliability of almost wholly unreflective, youthful intuitions as a premise for justifying her actions.

Much of this brings us back to the discussion of the internalised other in chapter four. Many of Briony's responses are ones which I would claim are a product of her emotional education. But what the description of the process above indicates both the complexity and flexibility of the various parts of our identity involved in such cases. The protectiveness towards the family is in conflict with her acquired attachment to honesty at an early stage in the narrative, and the conflict is resolved in favour of what she initially 'knows' to be right - the incrimination of Robbie. The changes which take place involve changes in her understanding of events, but also in her identity - humility, shame and the desire for atonement have now emerged as dominant aspects of who she is. These are transformations brought on by the interplay of the original features of the internalised other and the force of reason in throwing a new light on who she is and the actions for which she is responsible. Her perception of herself as having done the morally right thing is altered in light of the facts pressing in on her, and this in turns forces a change in her understanding of what kind of person she is. This process of change is only possible given the initial concern for others which was always a part of her through the presence of the internalised other, and the responsiveness of the internalised other to the light of reason.

If this sort of explanation is right, we need to consider a second concept of 'getting it right' which focuses not on propositional knowledge, but on the idea of wisdom. One advantage of such a concept is that it helps to explain why it is that time is needed for a someone to achieve understanding in ethics in a way in which it isn't in, say, mathematics. In this context, Raimond Gaita writes, 'In moral matters ...the achievement of deeper understanding requires that we have the depth to receive it, and that depth in ourselves is not a depository of propositions....but an historically *achieved* individuality.'[32] It is this 'deeper' understanding which only age can bring, and

[32] Raimond Gaita, *Good and Evil: An Absolute Conception*, 271.

without which our judgements are likely to go astray,[33] and it is for this reason that characteristics such as impetuousness, arrogance and impatience which are associated with youth constitute a persistent danger to sound judgement. Only age can give us a fuller awareness of our own finitude and the enormous complexity of what we are often confronted with. It is only from a more mature standpoint that we can reflect upon the very intuitions and emotions which guide us in the first place. It is when we do this that we are no longer simply acting, but acting in a way towards which we have a considered attitude. We not only act compassionately, but see it as a compassionate act, as a good act, an act which makes sense given the nature of who we are and of those around us. A child simply could not have this sort of more rounded understanding.

The thinker who has perhaps been most influential on this topic is Aristotle, and I shall therefore turn to him, and to the outline and paraphrase offered by David Wiggins, as well as Miles Burnyeat's interpretation of his discussion of the emotions.[34] Firstly, let us be clear about the terrain we are on. The key points at which sound judgement will be required will be when we are confronted with new challenges in the moral realm, where we have no precedents or bank of strictly relevant examples to fall back on. We are required to judge what to do in a situation where that judgement will represent a new step, and this means we must be clear about the unique features that each new situation will have, and the novel demands it will make on us. As Wiggins puts it, 'few situations come already inscribed with the names of all the concerns that they touch or impinge upon'.[35] It is rare that all the crucial features of the situation will be evident to us, and we require the ability to reflect and imagine in such a way that we take into account the full import of our possible actions. We may also find ourselves having to confront inconsistencies in our own outlook, torn perhaps between achieving what we are convinced is the right outcome, but reluctant to use what we consider questionable means. We may find ourselves forced to reconsider some of our deepest convictions on the grounds that they now stand in conflict with a cherished aim. We may find it is necessary to stop and start again, in that all the options hitherto considered are unpalatable. There must also be an awareness of ourselves as 'finite creatures who face an indefinite or infinite range of contingencies with only finite powers of prediction and imagination, of practical rationality itself.'[36]

It is in light of these sorts of considerations that we can see why wisdom could not be acquired in youth. Experience is the indispensable forerunner of wisdom because it is only when one has

---

[33] Cf Miles Burnyeat, 'The learner is envisaged as a young person who lives by the feelings of the moment and for that reason makes mistakes', 'Aristotle on Learning to by Good', reprinted in *Essays on Aristotle's Ethics*, ed A E Rorty, 78.
[34] David Wiggins, 'Deliberation and Practical Reason', reprinted in his *Needs, Values, Truth*, 215-238, and 'Aristotle on Learning to by Good' by Miles Burnyeat, reprinted in *Essays on Aristotle's Ethics*, 69-92,ed A E Rorty.
[35] David Wiggins, 'Deliberation and Practical Reason', reprinted in his *Needs, Values, Truth*, 231.
[36] Ibid, 233.

acquired a certain level of experience that the results of this can aid our judgements during the course of our practical reasoning. As Wiggins puts it, 'The person of real practical wisdom is the one who brings to bear upon a situation the greatest number of genuinely pertinent concerns and genuinely relevant considerations commensurate with the importance of the deliberative context.'[37] One must add that this sort of wisdom is not simply a question of accumulating lots of experiences, and it could only come over time. As such, it cannot be directly compared to acquiring new items of propositional knowledge. This in turn is explained by the fact that it is our emotions rather than our powers of cognitive reason which are at the heart of how we are learning. The emotions are shaped over time, not in an instant. The process of understanding is one which often requires a slow and painful relinquishing of previously held commitments, a process which we may resist. The conclusions we reach may be unwelcome, and we may be reluctant initially to contemplate the consequences. The key point is that the experiencing over time of our own mistakes, limitations and successes partly constitutes the acquisition of wisdom, and the kind of beings we are means that this can only be accomplished over time. The understanding of the wise agent is that of someone whose is familiar with her limitations at an experiential level, and this means we simply could not have a child prodigy in wisdom in the way we seem them in mathematics.

To return once more the example of Briony, it is the failure to understand the limited extent of her own powers of understanding as I have used the term in this context, and the inability to bring to bear a sufficiently wide range of relevant concerns and considerations which allow her to follow such a disastrous course of action despite her being motivated by her compassionate concerns for those around her. If one accepts this, then our tendency to defer to those who are 'older and wiser' makes sense as more than simply an adage, and Briony's change of heart reflects the changes in someone who has been through exactly the sort of painful process described in the previous paragraph, and is wiser as a result.

So far, I have focussed on the formal aspects of practical wisdom, the acquisition of the necessary skills, as it were, but what of the content? Aristotle famously claimed that only those with the right sort of emotional education could possibly benefit from his teaching,[38] but what will be the content of the views which shape the thoughts of the virtuous agent? We know that it is someone 'who has learned what is noble ....and ...come to love it.'[39] It is the agent whose emotions have been shaped in the appropriate way, and who has benefited from the right sort of formal education who can come to have the sort of practical wisdom under discussion here. Such an agent will have the right sort of dispositions, will come to enjoy acting in a fine and noble manner, will have an

---

[37] Ibid. 233.
[38] See Aristotle. Nicomachean Ethics. 1179b 5-19.
[39] Miles Burnyeat. 'Aristotle on Learning to be Good', reprinted in *Essays on Aristotle's Ethics*. ed A E Rorty. 76.

understanding of why she acts in the way she does, and will have acquired a level of experience which allows her to apply such skills in the face of novel and challenging circumstances. The content of such understanding will also be shaped by the nature of other persons, certain facts about their status as loving, suffering, caring beings.

But even this account of the wise and virtuous agent could not deliver up a rule along the lines of a calculation of utility which could given us an 'in principle' response to any conceivable situation. Our deliberations can aim at no more than applying the principles which constitute our reflective understanding of what is right in a novel situation, aware of our finite capacities, and informed by the facts of the matter. What we believe to be actions which are noble, just, fine, compassionate, heroic or ones which fall under any other such loaded descriptions cannot be given content independently of the ideals that an agent brings to the situation, and these ideals themselves will have a content which is shaped by what we see and discover and experience in our struggles in just such situations.

To be sure, there will be easy cases such as where we must choose between alleviating enormous suffering and giving up what we see as trivial sources of personal satisfaction. In such cases the virtuous agent will find it very easy to decide on which course of action will better reflect her moral values, but this falls short of some straightforward principle which can guarantee us something like an objectively right answer on all occasions. The most difficult circumstances we face are precisely those where we cannot weigh so easily the merits and demerits of what may be a wide range of possible actions and likely consequences, where we must deliberate in a patient, imaginative way to come to what we know will be no more than the best judgement we are capable of from our finite perspective. In such cases, we cannot start with a question such as what the compassionate response would be and then arrive by means of calculation at a correct answer, because it is in precisely such cases that we cannot know from the outset what the range of outcomes will be, how much suffering it may cause, how heavily the different sorts of suffering (eg the mental anguish vs mild physical pain) should weigh. Wisdom cannot guarantee us the kind of objectively right answers which we may wish for, merely the best ones we are capable of, and if this is a disappointing conclusion then it does no more than reflect the nature of the problem.

I am now in a position to return to the question of means versus ends, and to clarify their relationship in light of the comments of the last few pages. Firstly, neither can be seen as fully independent of the other. What is to count as an acceptable means must be conditioned by what we take to be our considered ends. We are not talking here about what Aurel Kolnai called the

'technical preamble' to our deliberation,[40] meaning certain relevant factual information such as numerical and causal connections. If we wish to help someone, the dispositions which provide this motivation in the first place will help to rule out certain courses of action. The agent who sees alleviating the suffering of others as an end is unlikely to see the causing of suffering as an attractive means. Even if this proves a necessary means, this in itself would be a subject for further deliberation on the part of the agent. The parent who holds her child as the doctor inoculates it sees the suffering as justified only because it avoids the danger of even greater suffering. Without that thought her actions would be unintelligible as those of a caring parent.

Similarly, our use of certain means may well alter what we take to be our ends. As Kolnai points out, 'Our ends are not ready-made, awaiting their fulfilment when the proper means should have been found; they may come to life and harden into shape in fairly unexpected contexts and their fixation involves to some extent .....a revision, modification and reorientation of our pre-established structure of permanent or comparatively lasting ends -I would rather say our *concerns* - itself.'[41] Within the context of what I have said earlier, this means that means and ends will have an essential interdependence, with each shaping, altering and influencing the other. Our ends in any given situation will proscribe certain courses of action and encourage others. The means may well be so daunting that they bring into question the value of the end, or else so costly in their use that they transform our view of the end to which they contribute. In all of these ways, each concept must be seen as part of a wider picture in constant flux, where our perception of specific means and ends are conditioned by the wider set of ideals and values we bring to situation, and where the onset of experience will change the ends we pursue and the means we find it acceptable to use, constrained and guided by certain features we find in ourselves and other human beings.

Reasoning with Others

I have focussed at great length on the form and content of the deliberations of the individual agent, but this form of personal deliberation is only part of the kind of moral reasoning we often engage in. One other key form such reasoning takes is when we try to persuade others of our moral point of view, often with the aim of getting them to act in a particular way, and I now wish to apply the principles of the previous sections of this chapter to this question.

We might begin by asking whether or not we really do reason with others in the moral domain, and it is a useful point of entry into the discussion to consider perhaps the most sceptical position on this issue, that of A J Ayer, who argued that 'one never really does dispute about questions of

---

[40] Aurel Kolnai, 'Delberation is of Ends', *Proceedings of the Aristotelian Society 62*, 200.
[41] Ibid, 205.

value.'[42] His reasons for holding this view were that moral words simply assert our own feelings on a particular matter and aim at exciting similar ones in the person with whom we are talking. As such, we do not assert anything about the particular object or action in question, but rather about our own attitude towards it with the aim of promoting the same attitude in the other person. Ethical judgements are 'mere' expressions of feeling expressed through a series of 'pseudo-concepts' which are ultimately meaningless, and are therefore worthy only of being jettisoned along with our language of theology and metaphysics. According to Ayer, a common process of moral conditioning is required for two agents to agree or disagree, and where this is absent we will tend to abandon any attempt at persuasion on the grounds that the other person 'has a distorted or undeveloped moral sense; which signifies merely that he employs a different set of values from our own.'[43]

Once one sifts out the derisory attitude towards ethics, there are nevertheless certain points of overlap between the position developed in this thesis and Ayer's view. I agree that our ethical language does express feelings we have towards a certain object, and I shall go on to say that when we argue we are trying to bring others to hold the same 'feeling-laden' beliefs that we hold. My arguments with regard to the importance of our social environment in the development of our emotions are also consistent with Ayer's discussion of 'conditioning'. From a taxonomical point of view, I have also been defending what could be described as a form of subjectivism. Where there are differences between my own position and Ayer's is that I think there is the possibility of moral dispute and moral reasoning with others.

To begin with, consider Ayer's analysis of our moral language. He argues that 'in saying that any type of action is right or wrong, I am not making any factual statement, not even as a statement about my own state of mind. I am merely expressing certain moral sentiments. And the man who is ostensibly contradicting me is merely expressing his moral sentiments, so there is plainly no sense in asking which of us is in the right in that neither of us is asserting a genuine proposition.'[44] The line of thought here is that the feelings which are expressed can be uncoupled from the propositions to which they are hooked up. Once we accomplish this, we can now translate the original propositions which have the form of factual descriptions about the world, into statements which describe what we are feeling. Statements such as 'X is good' and 'X is bad' can be translated into 'I like X' and 'I don't like X', and whereas we seemed to have a disagreement between two agents who had expressed different *beliefs* about the way the world is, we can now see that it is simply that they don't *feel* the same way about the way the world is.

---

[42] A J Ayer, *Language Truth and Logic*, 114.
[43] Ibid. 115.
[44] Ibid. 111.

The point of departure between Ayer's arguments and my own lies in his analysis of how our feelings relate to the propositions through which they are expressed. Rather than the propositions standing separately from the feelings, it is better to see the feelings as seared into the beliefs about the persons, types of action and categories of value to which they apply. If I believe that the death penalty is wrong, then we can go through the motions of talking about the modalities involved or talk of possible worlds in which the death penalty is right, but this just doesn't address the phenomenology of the way in which such beliefs are held. When one does hold such feeling-laden beliefs, they are held in such a way that the feeling and the proposition are held as inseparable, and it is only when we accept this that we can begin to understand why it is that people will die for such beliefs.[45] In light of this, consider the earlier passage by Styron quoted above. There is something absurd about claiming that the Nazi doctor and Sophie don't really disagree about how she and her children ought to be treated, and the reason it appears so absurd is because such a theory ignores the very real way in which the feelings and beliefs involved in such thought just cannot be separated out from one another in the forensic way suggested in *Language, Truth and Logic*.

In light of this, I suggest Ayer's claim that there could be no moral beliefs needs substantial revision. I accept that moral beliefs do not express facts about the world in the way a scientific theory aims to, but the way to address this is not to separate out the feeling and the propositional content and to dispense with the whole domain as meaningless, but to recognise that moral beliefs are a distinctive category of belief, distinctive because of their feeling-laden nature, and capable of forming the source of a distinctive form of reasoning with others - moral reasoning.

Let me now turn to Ayer's talk of the importance of what he calls conditioning. I have argued throughout that our emotions are indeed shaped by our social environment,[46] and that our moral beliefs, in some sense, bottom out in our emotional and feelingful responses to the world around us. But the theory of emotions I have defended and the discussion in the previous chapter provide the basis for understanding how moral views can vary in a way which conforms to certain moral intuitions and to our observations of moral reasoning.

In the first instance, we must distinguish between the claims that our emotions (and hence our moral perspective) are shaped by the social environment, and the claim that they are entirely

---

[45] I accept that there are cases where such beliefs might 'go dead' on us, and where we might continue to express them more through habit than because they retain the feeling-laden force they once had. In this respect, I accept the contingent relationship between the feeling and the proposition, but dispute that this sort of separation is one which is consistent with the phenomenology of such beliefs across the board.

[46] See in particular 56-62 above.

determined by it. The empirical research canvassed in chapter two[47] indicated that certain features of our emotional lives appear to be cross-cultural. That is to say that the expression of emotions and our capacity to grasp the emotions others are experiencing have a common core which even those from cultures who have had no contact with one another can discern. This supports the more colourful claim of William James that if you see an elephant charging towards you then your cultural background will probably have very little influence on your reaction.

This in turn lends support to the 'sensible subjectivism' of Wiggins, in which we should understand the interplay between the subject and the world as one in which 'there is something in the object that is *made for* the sentiment it would occasion in a qualified judge'.[48] To take the example of compassion, then the suffering we observe brings down the sentiment upon it, a sentiment '*made for*' this sort of event. Once we have certain features of the world which elicit emotional responses that are universal in mankind, then we have the means to restrict any form of relativism which may emerge as a result of acknowledging the developmentally open nature of our emotions, and at the same time explain why a divergence in the full content of our moral perspectives is to be expected within and across different cultures given that developmentally open nature.

In order to elucidate this, let me offer an outline of how this sort of theory would describe what we see in practice. If we imagine a person born into a particular community, then she starts out with certain relevant capacities which are universal, including the capacity to grasp the responses of others and the ability to have her feelings and emotions shaped according to those responses. Other general features such as aversion to pain make it likely that there will be a high degree of convergence across different communities with regard to certain responses the agent will elicit or witness around her, and these responses will result in the internalised other taking on a content which is likely to involve the kinds of ideals which are likely to overlap with those of agents brought up in different communities, given the kind of creatures we are. Once we take into account these crucial common characteristics, then we have the basis for discussing them in the context of a wider moral theory, it is surely for this reason that characteristics such as our aversion to pain and our attraction to pleasure are capable of playing such a central role in the thought of both Hume and Aristotle.

The outcome of this is that we can engage in discussion capable of producing a decisive conclusion insofar as there is a consistency between agents whose feelings have come to influence a similar range of beliefs in a similar way. Within a single community there is likely to be a greater degree of convergence, less so across different communities. But there is always, in principle, the likelihood

---

[47] See page 47 above.
[48] David Wiggins, 'A Sensible Subjectivism', reprinted in his *Needs, Values, Truth*, 194.

that our common humanity will generate more rather than less potential for discussion and persuasion. A mother's loving care for her child is likely to enter the scheme of values in any culture, and this means that the ability of one mother to respond to the suffering of another who has lost her child is created, as is the possibility of a wider moral discussion on this subject. If this is correct, a sentimentalist theory of ethics retains the possibility of universal standards even if we accept the developmentally open nature of our emotions. Once there are shared feeling-laden beliefs, then there is the opportunity for reasoning and discussing with others which can result in a clear agreement. We will search for agreement because we are pressed into moral discussion by the feelings and emotions which condition our responses to the world around us, and guide our thinking. Our aim in individual deliberation is to settle on the right course of action, and our aim in reasoning with others will be to bring them to the point where they see things aright. To do so, we appeal to a shared set of feeling-laden beliefs about the world, and reason from these premises to a conclusion which we can both 'see' in light of these premises. In this context, coming to see things aright will mean acquiring a similar understanding and set of responses to our interlocutor, and thereby coming to agree on which course of action would be the best.

The drawback is that such an outline also throws up at least one major limitation on what could possibly be achieved through a process of reasoning with others. For all the universality of our physical make-up, our dispositions to love our children, our aversion to pain, these factors simply make morality more likely to develop in some ways rather than others. In the final analysis, the precise content of any specific moral belief will be contingent upon the vagaries of the community within which the agent is raised, and this theory is consistent with the view that there may be points of fundamental moral importance within one community to which there is an equally fundamental opposition in another.

The moral realist holds out the hope that any such dispute might be settled by the moral fact of the matter, an option not open to the subjectivist. There is little option but to bite this bullet, and identify this as an unfortunate but unavoidable aspect of the human condition. In such circumstances we may be able to throw up our hands in frustration and agree to disagree. We may be able to reach some sort of compromise. But we may also find ourselves in the position where there is simply no alternative to a conflict which we will describe as a just war.[49] The conclusions of the previous section about the open-ended nature of moral deliberation also hold for reasoning with others, and on this theory it is not even possible in principle that there could be a final set of moral ideals which can be translated on any given occasion into an answer to a particular moral dispute. It simply turns out that there are going to be occasions where we cannot persuade our disputant of

---

[49] For a wide-ranging discussion of the different possible approaches to this question, see Alan Gibbard, *Wise Choices, Apt Feelings*, Chapter Thirteen.

our own position, and our position is such that we cannot compromise on it. Not only this, but we will find ourselves in a position which may well be characterised by a form of incomprehension. The position of our disputant will be one which we simply cannot understand given our own outlook.

There is a powerful and disturbing exploration of this in J M Coetzee's novel *Waiting for the Barbarians*, in which a former magistrate who has fallen foul of his country's tyrannical regime tries to comprehend the viewpoint of his torturer.

'I am trying very hard to understand your feelings towards me.....Much more than an opportunity to address these people to whom I have nothing to say, would I appreciate a few words from you. So that I can come to understand why you devote yourself to this work. And can hear what you feel towards me, whom you have hurt a great deal and now seem to be proposing to kill.'

'Do you see this hand?' he says. He holds his hand an inch from my face. 'When I was younger' - he flexes the fingers - 'I used to be able to poke this finger' - he holds up the index finger - 'through a pumpkin-shell.' He puts the tip of his finger against my forehead and presses. I take a step backwards.[50]

The response, of course, is chilling, but there is nothing to suggest that either of the two has any understanding of what it is like to view the world in the way the other does, and this is confirmed later when the victim sincerely appeals once again for answers to the same sorts of questions.

> 'Forgive me if the question seems impudent, but I would like to ask: How do you find it possible to eat afterwards, after you have been .....working with people? That is a question I have always asked myself about executioners and other such people. Wait! Listen to me a moment longer, I am sincere, it has cost me a great deal to come out this, since I am terrified of you, I need not tell you that, I am sure you are aware of it. Do you find it easy to take food afterwards? I have imagined that one would want to wash one's hands. But no ordinary washing would be enough, one would requires priestly intervention, a ceremonial of cleansing, don't you think? some kind of purging of one's soul too - that is how I have imagined it. Otherwise how would it be possible to return to everyday life - to sit down at a table for instance, and break bread with one's family or comrades?'.....I am trying to

[50] J M Coetzee, *Waiting for the Barbarians*, 129.

> understand the zone in which you live. I am trying to imagine how you breathe and eat and live from day to day. But I cannot! That is what troubles me![51]

This time he is met not with the detached, cold nostalgia of the previous occasion, but with frightened abuse from a figure who has now come under threat himself as events move against him. But what is illuminating here is the genuine failure to understand the way in which someone else can conduct his life in a fashion so utterly devoid of the values and dispositions which shape one's own. What we see hear is not simply disagreement, but bemusement, and this stems precisely from a set of emotional responses which provide no common ground for understanding. The beliefs that each man has with regard to torture and execution lack a comparable feeling-laden element, thereby making understanding between the two impossible. It is when such instances take on a more generalised form that agreement on moral questions could not occur, because there is no pre-existing common emotional framework against which beliefs are formed. Such disputes could not result in agreement, because there could be no understanding out of which it might emerge, no consistency in the content of our moral concepts. It is here where we find the most depressing aspect of the theory set out here - the denial of the prospect of universal moral agreement around a set of objectively correct moral principles.

Simple Moral Responses

The accounts of moral reasoning which I have discussed so far have focussed on processes which involve what one might think of a process which aims to answer questions such as 'What is the right thing to do?', where we are required to consider at some length the appropriate course of action, and where this may not be at all obvious to us from the outset. But something now needs to be said about the sorts of simpler responses which we often see, where there is no obvious process of deliberation. In circumstances where we see the person is in need of help, and just help without thinking about it, what is going on?

Much of what has been said in the previous three chapters will help to explain this sort of response. Firstly, the perception of need must be one which is salient in the cognitive economy of the agent. John McDowell makes such a point when he claims that 'The sensitivity is, we might say, a sort of perceptual capacity.'[52] This means that the person who responds compassionately when seeing someone is suffering is not necessarily thinking to herself that she must now offer a compassionate response, but is rather someone who is notices suffering when she is confronted with it, sees is as something to be addressed, and is disposed to help where she can. The virtuous agent is someone

---

[51] *Ibid.* 138.
[52] John McDowell, 'Virtue and Reason', reprinted in *Mind, Value and Reality*.

who will, as it were, just see helping as 'the thing to do'. McDowell once again captures the sort of underlying feature of the virtuous agent in such circumstances when he describes virtue in general as 'an ability to recognise requirements that situations impose on one's behaviour. It is a single complex sensitivity of this sort we are aiming to instil when we aim to inculcate a moral outlook'.[53]

When we say that the virtuous agent has the sort of sensitivity to the situations she encounters which will lead her to simply act without need of lengthy deliberation, this does not mean that she could not offer an explanation if required. The virtuous agent may well respond to questions such as 'Why did you help her?' with general answers such as 'Because she needed it' or 'Because it is was the thing to do', but this does not make it akin to a blind, instinctive response. The dispositions of the virtuous agent are revealed in her sensitivity to certain situations and her tendency to respond in certain ways, but this also reveals a much wider web of beliefs one would expect such an agent to have. These will include considerations such as how to conduct oneself with others, how to treat others, and will also reveal deeper beliefs about the kind of life a human being should lead. It is in the context of these sorts of beliefs that individual actions and response will make sense.

> If someone guides his life by a certain conception of how to live, then he acts on particular occasions, so as to fulfil suitable concerns. A concern can mesh with a noticed fact about a situation, so as to account of an action; as, for instance, a concern for the welfare of one's friends, together with awareness that a friend is in trouble and open to being comforted, can explain missing a pleasant party in order to talk to the friend.[54]

As such, the actions of the agent, when performed consistently, reveal the kind of sensitivities the agent has, and thereby the kind of wider values which inform the agent's life. The arguments advanced over the previous chapters of this thesis indicate both how the virtuous agent comes to have the sort of responses she has, and what responses will count as virtuous. Acquiring the sensitivities of the virtuous agent is a function of the education of the emotions. This agent reacts in the way she does because her feelings have been shaped such that certain features of the world now become more salient than others when she perceives them, and this salience is a feeling-laden aspect to the beliefs and perceptions of the world we find before us. It is this feature of our identity which I have described as the internalised other.

The content of the responses of the virtuous agent - the agent who has a proper moral perspective - is a content which matches up to our need for others, the ability of others to respond to those needs, and our ability to respond to their needs. Needs are revealed through suffering, and the

---

[53] *Ibid*, 53.
[54] *Ibid*, 67.

response of the virtuous agent to suffering is compassion. This means the kind of simple responses where we see someone acts without the kind of deliberation or discussion described in the previous two sections are the responses of someone who has been educated to recognise and respond to the needs of others, and is therefore disposed to help others when such circumstances present themselves.

Conclusion

By way of conclusion, I shall set out in broad outline the consequences this theory has for metaethics. More specifically, I shall suggest the direction and limitations of any conclusions in ethics given the nature of this thesis.

The first observation one can make is that the ethical domain is one which claims us long before we come to reflect on it. The internalised other is formed from a very early stage, and the nature of our emotional attachments which stem from this source ensure that we are bound within the moral sphere. One property of emotions is that they have a passive element - there is an involuntary side to any emotional response, and if emotions play a crucial role within the ethical domain it means we are involuntarily drawn into that domain by the nature of our responses. This is what I mean by the ethical claiming us. It also follows that there could be no 'view from nowhere' with regard to the ethical, no neutral position from which we could determine what the best content for morality will be. We are all raised within communities which shape our emotions in such a way that we come to understand the world in ways conditioned by that social education. This means that all moral reflection takes place from a moral perspective. To be sure, we can theorise about morality, commenting on the behavioural components, the cross-cultural comparisons, the interest it has provoked within philosophy, but the moment we address specifically moral questions such as how we ought to treat others then an answer could only be arrived at under the guidance of the feelings and emotions which give morality content in the first place. To be without these feelings and emotions would indeed give us a perspective from the moral nowhere, but such a position would be one from where the very question would have neither the force nor the content morality entails. It is only when I have some sort of emotional engagement with others around me that life is breathed into morality. In that sense, it is closer to love than it is to reason.

A further consequence is that there could be no final moral answers, where this would be an objective account of what morality demands of us. There will certainly be cases where all of the wisdom and knowledge we have indicates that particular courses of action are right or wrong, but my point is that this conviction could never stem from an objective account of what right and wrong are. Our morality is partly constituted by our responses to the world around us, and our

sentiments are brought down upon the world by what we find there. But the complexity of the human condition means that morality must always remain open-ended. It is an inescapable aspect of human life that we are constantly confronted with novel circumstances which require acts of judgement as to what morality requires, and where there is no detached, right answer which could be available to us. What counts as right in such cases will be partially determined by what we see and how we feel, and even if we managed to reach some unlikely point of universal consensus, it would still be no more than a contingent judgement on how to act rather then a correct response objectively correlated to the way the world stands.

In light of this, what can we hope to achieve through deliberation about morality? We can begin by acknowledging our own status as moral agents - people already claimed by ethics and who are already engaged in the world with a perspective. We can also recognise the necessity and the limitations of what we are engaged in. Necessity, given the partly involuntary nature of our responses, and limitations in the sense that there could be no final stamp of approval for the judgements we arrive at. David Wiggins makes both these points in setting out his sensible subjectivism.

> In this matter the subjectivist really has to do the same as everyone else; he can only urge that, in spite of the possibility of irresolvable substantive disagreement, but in a manner partially conditioned by that possibility, we should persevere as best we can in the familiar process of reasoning, conversion, and criticisms - without guarantees of success, which are almost as needless as they are unobtainable.[55]

We can do our best to ensure that we are aware of all the relevant facts, and that we are reasoning soundly. We can strive for a life which is consistent with what our best account of how we ought to live, and we can try to remain aware of the limitations of our abilities. For anyone who believes that we are required to produce much more than this for a satisfactory moral theory this will seem a pallid attempt at understanding what we could achieve. Opposition would doubtless centre on the view that our emotions are so crucial to our morality, and might well look to restore a more prominent position for reason over what may be seen as the 'slavish' way in which emotion directs our behaviour. There is no more I can say in defence of my position than has appeared on the pages of this thesis.

---

[55] David Wiggins, 'A Sensible Subjectivism', reprinted in his *Needs, Values, Truth*, 210.

Bibliography

Aristotle *Nicomachean Ethics*. c. 325 BC. Tr. T. Irwin. Indianapolis: Hackett, 1985.

Rhetoric. Tr. W Rhys Roberts. Reprinted in Barnes (1984).

Anscombe, E. 'Modern Moral Philosophy'. *Philosophy* 33, 1958, 1-19.

Armon-Jones, C. 'The Thesis of Constructionism'. In Harre (1986), 32-56.

Armstrong, D. *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul, 1968.

Ayer, A. J. *Language, Truth and Logic*. London: Gollancz, 1946.

Barnes, J. (ed.), *The Complete Works of Aristotle: the Revised Oxford Translation*, 2 Vols, Princeton N.J.: Princeton University Press, 1984.

Barnett, M. A. 'Empathy and Related Responses in Children'. In Eisenberg and Strayer (1987), 146-62.

Bennett, J. *Linguistic Behaviour*. Indianapolis: Hackett, 1990.

Berlin, I. *Four Essays on Liberty*. Oxford: Oxford University Press, 1969.

Bermudez, J. L., Marcel, A, Eilan, N. *The Body and the Self*. Cambridge Mass.: MIT Press, 1998.

Blackburn, S. 'How To Be an Ethical Antirealist'. In *Midwest Studies in Philosophy* 12, 1988, 361-75

Ruling Passions: A Theory of Practical Reasoning. Oxford: Clarendon Press, 1998.

'How Emotional is the Virtuous Person?' In Goldie (2002b), 81-96.

Blum, L. *Moral Perception and Particularity*. Cambridge: Cambridge University Press, 1994.

Bond, E. J. *Reason and Value*. Cambridge: Cambridge University Press, 1983.

Brewer, B. 'Emotion and Other Minds'. In Goldie (2002b), 23-36.

Budd, M. *Wittgenstein's Philosophy of Psychology*. London: Routledge, 1989.

Burnyeat, M. 'Aristotle on Learning to be Good'. In A . O. Rorty (1980b), 69-92.

Butler J, *Fifteen Sermons Preached at the Rolls Chapel.* In D. D. Raphael (1969), Vol,. 1, 325-377.

Coetzee, J. M. *Waiting for the Barbarians.* London: Vintage, 2000 edn.

Chalmers, D. *The Conscious Mind.* Oxford: Oxford University Press, 1996.

Cohen, D., Vandello, J, Rantilla, A. K. 'The Sacred and the Social: Cultures of Honor and Violence'. In Andrews and Gilbert (1998), 261-282.

Damasio, A. R. *Descartes' Error.* London: Macmillan, 1994.

Dancy, J. *Moral Reasons.* Oxford: Blackwell, 1993.

Daniels, N. (ed.), *Reading Rawls.* Oxford: Blackwell, 1975.

D'Arms, J., Jacobson, D., 'Expressivism, Morality and the Emotions'. *Ethics* 104, 1994, 739-763.

Darwall, S., Gibbard, A., Railton, P. 'Toward *Fin de Siecle* Ethics: Some Trends'. *Philosophical Review* 1992: 115-89.

Darwin, C. *The Expression of Emotion in Man and Animals, with an introduction, afterword and commentary by P. Ekman.* London: Harper Collins, 1998.

Davidson, D. 'Actions, Reasons and Causes'. In *Journal of Philosophy* 60, (1963).
'Mental Events'. Reprinted in Davidson (1980), 207-225.
'Hume's Cognitive Theory of Pride. Reprinted in Davidson (1980), 277-290.
Essays on Actions and Events. Oxford: Oxford University Press, 1980.
'The Irreducibility of the Concept of the Self'. Reprinted in Davidson (2001), 85-93.
'Rational Animals'. Reprinted in Davidson (2001), 95-107.
'The Second Person'. Reprinted in Davidson (2001), 107-123.
'The Emergence of Thought'. Reprinted in Davidson (2001), 123-135.
Subjective, Intersubjective, Objective. Oxford: Clarendon Press, 2001.

Deigh, J. 'Cognitivism in Theory of the Emotions'. *Ethics* 104, 824-54.

Dennett, D. *Kinds of Minds: The Origins of Consciousness.* London: Phoenix, 1996.

Descartes, R. *Passions of the Soul*, in *The Philosophical Writings of Descartes.* Tr. J. Cottingahm, R. Stoothoff, and D. Murdoch. Vol. 1, Cambridge: Cambridge University Press, 1985, 325-404.

Dworkin, R. *Sovereign Virtue: The Theory and Practice of Equality.* Cambridge Mass.: Harvard University Press, 2000.

Dunn, J., Brown, J., Beardsall, L. 'Family Talk about Feeling States and Children's Later Understanding of Other's Emotions'. In Jenkins, Oatley and Stein (1998), 168-183.

Eisenberg, N., Strayer, J. (eds) *Empathy and its Development.* Cambridge: Cambridge University Press, 1987.

Ekman, P. and Friesen, W. V. 'Constants across Culture in the Face of Emotion'. Reprinted in Jenkins, Oatley and Stein (1998), 63-73.

Foot, P. *Virtues and Vices.* Oxford: Clarendon Press, 2002 edn.
Natural Goodness. Oxford: Clarendon Press, 2001.

Franzen, J. *The Corrections.* New York: Fourth Estate, 2001.

Gaita, R. *Good and Evil: An Absolute Conception.* Basingstoke: Macmillan, 1991.

Gibbard, A. *Wise Choices, Apt Feelings: a Theory of Normative Judgement.* Oxford: Clarendon Press, 1990.

Gilbert, P., Andrews, A. (eds) *Shame: Interpersonal Behaviour, Psychopathology, and Culture.* Oxford: Oxford University Press, 1998.

Goldie, P. *The Emotions: A Philosophical Exploration.* Oxford: Oxford University Press, 2000.
'Can We Trust Our Emotions?'. In *Richmond Journal of Philosophy* Vol. 1, Issue 1, Summer 2002(a), 27-36.
Understanding Emotions (ed.). Aldershot: Ashgate, 2002(b).

'Narrative and Perspective; Values and Appropriate Emotions'. In A. Hatzimoysis (ed.) *Philosophy and the Emotions*. Cambridge: Cambridge University Press, 2003, 201-220.

On Personality. London: Routledge, 2004

Griffiths, P. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: Chicago University Press, 1997.

Guyer, P. (ed.) *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press, 1992.

Hamilton, E., Cairns, H. (eds) *The Collected Dialogues of Plato, including the letters*. Princeton N.J.: Princeton University Press, 1971.

Hare, R. M. *The Language of Morals*. Oxford: Oxford University Press, 1952.

Harman, G. *The Nature of Morality: An Introduction to Ethics*. Oxford: Oxford University Press, 1977.

Harre, R. (ed.) *The Social Construction of the Emotions*. Oxford: Blackwell, 1986.

Harris, P. *Children and Emotion: The Development of Psychological Understanding*. Oxford: Blackwell, 1989.

Hatzimoysis, A. 'Emotional Feelings and Intentionalism'. In A. Hatzimoysis (ed.) *Philosophy and the Emotions*. Cambridge: Cambridge University Press, 2003, 105-112.

Hobbes, T. *Leviathan*. Ed. C. B. Macpherson. London: Penguin, 1968.

Holy, L. *The Little Czech and the Great Czech Nation*. Cambridge: Cambridge University Press, 1996.

Homer *The Iliad*. Tr. Robert Fagles. New York: Viking Penguin, 1990.

Hume, D. *A Treatise of Human Nature*. 1739-40, ed. L. A. Selby-Bigge. Oxford: Oxford University Press, 1978.

An Enquiry concerning the Principles of Morals, 1748, ed. J. B. Schneewind. Indianapolis: Hackett, 1983

Hursthouse, R. *On Virtue Ethics*. Oxford: Oxford University Press, 1999.

Hutto. D. 'The World is not Enough: Shared Emotions and Other Minds'. In Goldie (2002b), 37-54.

James, S. *Passion and Action.* Oxford: Clarendon Press, 1997.

James, W. 'What is an Emotion?'. In *Mind* 9, 188-205.

Kant, I. *Groundwork of the Metaphysics of Morals.* Tr. H. J. Paton. New York: Harper Torchbooks, 1964.

Critique of Practical Reason. Tr. Mary Gregor. Cambridge: Cambridge University Press, 1997.

Kenny, A.. *Action, Emotion and Will.* London: Routledge, 1963.

Kitto, H. D. F. *The Greeks.* London: Penguin, 1991 edn.

Kolnai, A. 'Deliberation is of Ends'. In *Proceedings of the Aristotelian Society* 62: 195-218.

Korsgaard, C. M. *Creating the Kingdom of Ends.* Cambridge: Cambridge University Press, 1996(a).

The Sources of Normativity: Cambridge: Cambridge University Press, 1996(b).

Kupperman, J. J. 'The Emotions of Altruism in East and West'. In Marks J. and Ames R. T. (eds.), *Emotions in Asian Thought: A Dialogue in Comparative Philosophy.* Albany: State University of New York Press, 1995, 123-38.

di Lampedusa, G. T. *The Leopard.* Tr. Archibald Colquoun. London: Harvill, 1966.

Lewis M, Sullivan M W, Stanger C, and Weiss M, 'Self Development and Self-Conscious Emotions'. In Jenkins, Oatley and Stein (1998).

Lindisfarne, N. 'Gender, Shame and Culture: An Anthropological Perspective'. In Gilbert and Andrews (1998), 246-260.

Mackie, J. *Ethics: Inventing Right and Wrong.* London: Penguin, 1977.

Hume's Moral Theory. London: Routledge, Kegan and Paul, 1980.

McDowell, J. 'Virtue and Reason'. In *Monist* 62, 331-50.

'Are Moral Requirements Hypothetical Imperatives'. In *Proceedings of the Aristotelian Society*, supp. vol. 52 (1978): 13-29.

'Values and Secondary Qualities'. Reprinted in Darwall, Gibbard and Railton (1997), 201-215

'Projection and Truth in Ethics'. Reprinted in Darwall, Gibbard and Railton (1997), 215-227.

Mind and World. Cambridge Mass.: Harvard University Press, 1994.

Mind, Value and Reality. Cambridge Mass: Harvard University Press, 1998.

MacIntyre, A. *After Virtue*. Notre Dame, Ind.: University of Notre Dame Press, 1981.

Main, M., George, C. 'Responses of Abused and Disadvantaged Toddlers to Distress in Agemates: A Study in the Day Care Setting'. In *Developmental Psychology* 21, 407-12.

Malcolm, N. 'Thoughtless Brutes'. In *Proceedings of the American Philosophical Association* 46, 1972-73.

Martin, M. G. F. 'Bodily Awareness: A Sense of Ownership'. In Bermudez, Marcel and Eilan (1998), 267-290.

McEwan, I. *Atonement*. London: Vintage, 2002.

Milgram, E. *Practical Induction*. Cambridge Mass.: Harvard University Press, 1997.

Miller, D. (ed.) *Liberty*. Oxford: Oxford University Press, 1991.

Moore, G. E. *Principia Ethica*. Cambridge: Cambridge University Press, 1903.

Morris, H. (ed.) *Guilt and Shame*. London: Wadsworth, 1971.

Morsbach, H. and Tyler, W. J. 'A Japanese Emotion: *amae*', in Harre (1986), 289-307.

Moya, C. J. *The Philosophy of Action: An Introduction*. Cambridge: Polity, 1990.

Murdoch, I. *The Sovereignty of Good*. London: Cambridge University Press, 1970.

Nagel, T. *The Possibility of Altruism*. Princeton: Princeton University Press, 1978.

Nietzsche, F. *Beyond Good and Evil: Prelude to a Philosophy of the Future.* Tr. R. J. Hollingdale. Harmondsworth: Penguin 1990.

On the Genealogy of Morals: a Polemic by way of clarification and supplement to my last book. Tr. Francis Golffing. New York: Doubleday, 1956.

Beyond Good and Evil. Tr. Douglas Smith. Oxford: Oxford University Press, 1996.

Nozick, R. *Anarchy, State and Utopia.* Oxford: Blackwell, 1974.

Nussbaum, M. *Upheavals of Thought: The Intelligence of Emotions.* Cambridge: Cambridge University Press, 2001.

Oakley, J. *Morality and the Emotions.* London: Routledge, 1992.

O'Hear, A. (ed.) *Current Issues in Philosophy of Mind.* Oxford: Oxford University Press, 1998.

Ondaajte, M. *The English Patient.* London: Picador, 1993.

O'Shaughnessy, B. *The Will: a Dual Aspect Theory.* Vol. 1. Cambridge: Cambridge University Press, 1980.

'Proprioception and the Body'. In Bermudez, Marcel and Eilan (1998), 175-204.

Papineau, D. *Thinking about Consciousness.* Oxford: Clarendon Press, 2002.

Pitcher, G. 'Emotion'. In *Mind* 74, 326-46.

Plato, *The Republic*, in Hamilton and Cairns (1961).

The Phaedrus, in Hamilton and Cairns (1961).

Quine, W. V. 'Two Dogmas of Empiricism'. In *Philosophical Review* 60, 1951, 20-43.

Ontological Relativity and Other Essays. New York: Columbia University Press, 1969.

Raphael, D. D. (ed.) *The British Moralists from Hobbes to Bentham.* 2 Vols. Oxford: Clarendon Press, 1969.

Rawls, J. *A Theory of Justice.* Oxford: Oxford University Press, 1973.

Political Liberalism. New York: Columbia University Press, 1993.

Justice as Fairness. Cambridge Mass.: Harvard University Press, 2001.

Rorty, A. O. (ed.) *Explaining Emotions*. Berkeley: University of California Press, 1980 (a).

(ed.) *Essays on Aristotle's Ethics*. Berkeley: University of California Press, 1980 (b).

'Jealousy, Attention and Loss' in Rorty (1980a).

Rorty, R. (ed.) *The Linguistic Turn: Essays in Philosophical Method With Two Retrospective Essays*. Chicago: University of Chicago Press, 1992.

Rosen, M., Wolff, J. (eds) *Political Thought*. Oxford: Oxford University Press, 1999.

Ryle, G. *The Concept of Mind*. London: Penguin, 1949.

Sartre, J-P. *Being and Nothingness*. Tr. H. E. Barnes. London: Routledge, 1943.

Sketch for a Theory of the Emotions. Tr. P. Mairet. London: Methuen, 1962.

Scanlon, T. M. 'Contractualism and Utilitarianism'. Reprinted in A. Sen and B. Williams (eds), Utilitarianism and Beyond, Cambridge: Cambridge University Press, 1982.

Schopenhauer, A. *On the Basis of Morality*. Tr. E. F. J. Payne. Providence RI.: Berhahn Books, 1995.

Schore, A. N. 'Early Shame Experiences and Infant Brain Development'. In Andrews and Gilbert (1998), 57-77.

Schueler, G. F. *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge Mass.: MIT Press, 1995.

Scruton, R. 'Emotion, Practical Knowledge and Common Culture'. In Rorty (1980a).

Smith, A. *A Theory of Moral Sentiments*. Reprinted in Raphael (1969), Vol. 2, 202-254.

Smith, M. *The Moral Problem*, Oxford: Blackwell, 1994.

"The Possibility of Philosophy of Action'. In J. Bransen (ed.), *Human Action, Deliberation and Causation*, Dordrecht, Netherlands: Kluwer Academic Publishers, 17-41.

Solomon, R. 'Emotions and Choice'. In A. O. Rorty (1980a), 251-282.

The Passions: Emotions and the Meaning of Life. Indianapolis: Hackett, 1993.

Stocker, M. 'Psychic Feelings: Their Importance and Irreducibility'. In Australian Journal of Philosophy, 61, 5-26.
with Hegeman, E., *Valuing Emotions*. Cambridge: Cambridge University Press, 1996.

Styron, W. *Sophie's Choice*. London: Bantam, 1979.

Taylor, C. *Sources of the Self: the Making of the Modern Identity*. Cambridge: Cambridge University Press, 1989.

Thompson, R. A. 'Empathy and Emotional Understanding: the Early Development of Empathy'. In Eisenberg and Strayer, 1987, 119-45.

Taylor, G. *Pride, Shame and Guilt: Emotions of Self-Assessment*. Oxford: Oxford University Press, 1985.

Tolstoy, L. *War and Peace*, Oxford World's Classics edn. Tr. L. and A. Maude. Oxford: Oxford University Press, 1933.
Anna Karenina, Oxford World's Classics edn. Tr. L. and A. Maude. Oxford: Oxford University Press, 1980.

Wiggins, D. 'Categorical Requirements: Kant and Hume on the Idea of a Duty'. In *The Monist*, Vol. 74, No. 1 (1991)
Needs, Values, Truth, 3rd edn. Oxford: Clarendon Press, 1998.
'Truth, Invention and the Meaning of Life'. In Wiggins (1998), 87-138.
'A Sensible Subjectivism?'. In Wiggins (1998), 185-214.
'Deliberation and Practical Reason'. In Wiggins (1998), 215-238.

Williams, B. 'Morality and the Emotions'. In Williams 1973, 26-45.
'Egoism and Altruism'. In Williams 1973, 250-265.
'The Idea of Equality'. In Williams 1973, 230-249.
Problems of the Self. Cambridge: Cambridge University Press, 1973.
'Internal and External Reasons'. In *Moral Luck*, Cambridge: Cambridge University Press, 1981, 101-113.
'Practical Necessity'. In *Moral Luck*, Cambridge: Cambridge University Press, 1981, 124-132.
Ethics and the Limits of Philosophy. London: Fontana, 1985.

Shame and Necessity. Berkeley: California University Press, 1993.

'Moral Incapacity', in Williams (1993), 46-56.

Wittgenstein, L. *Philosophical Investigations*. Oxford: Blackwell, 1958.

Wolff, J. *An Introduction to Political Philosophy*. Oxford: Oxford University Press, 1996.

Wollheim, R. *The Thread of Life*, Cambridge Mass.: Harvard University Press, 1984.

On the Emotions, New Haven, Conn.: Yale University Press, 1999.

Zahn-Waxler, C., Radke-Yarrow, M. 'The Development of Altruism: Alternative Research Strategies'. In Eisenberg-Berg N. (ed.), *The Development of Prosocial Behaviour*. New York: Academic Press, 1982.